DOCUMENT RESUME

ED 130 599 IR 004 127

AUTHOR	Silverman, Robert E.
TITLE	Automated Teaching: A Review of Theory and Research. Technical Report.
INSTITUTION	Naval Training Device Center, Orlando, Fla.
REPORT NO	NAVTRADEVCEN-507-2
PUB DATE	8 Jun 60
NOTE	54p.; Archival document
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	*Autoinstructional Aids; Branching; Comparative Analysis; Conventional Instruction; Literature Reviews; Motivation; *Programed Instruction; Programing; Prompting; Reinforcement; *Teaching Machines
IDENTIFIERS	Size of Step

ABSTRACT

The present state of research in the area of automated teaching and the application of automated teaching devices were reviewed in terms of issues relating to programing, to machine variables, and to studies comparing conventional instruction with automated instruction. Some basic issues were singled out for consideration. Such problems as prompting, confirmation, reinforcement, size-of-step, and branching were considered. The design and construction of machines were judged to be a secondary issue; the question of programing takes precedence. The problem of measuring the effects of automated teaching was considered to be crucial, since it is bound to the question of what is to be learned. It was also pointed out that job analyses of learning tasks must be done and measurement problems solved before valid comparisons between automated and conventional instruction can be made. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED130599

Technical Report: NAVTRADEVCEN 507-2

AUTOMATED TEACHING: A REVIEW OF THEORY AND RESEARCH

Prepared by:

Robert E. Silverman
Department of Psychology
New York University
Contract N61339-507

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Approved:

Edw C. Callahan

Edw. C. Callahan, Captain, USN
Commanding Officer and Director

Distribution:
Special List

U. S. NAVAL TRAINING DEVICE CENTER
PORT WASHINGTON, NEW YORK

8 June 1960

HR041107

ABSTRACT

AUTOMATED TEACHING: A REVIEW OF THEORY AND RESEARCH

The present state of research in the area of automated teaching and the application of automated teaching devices were reviewed in terms of issues relating to programming, to machine variables, and to studies comparing conventional instruction with automated instruction. Some basic issues were singled out for consideration. Such problems as prompting, confirmation, reinforcement, size-of-step, and branching were considered. The design and construction of machines were judged to be a secondary issue; the question of programming takes precedence. The problem of measuring the effects of automated teaching was considered to be crucial, since it is bound to the question of what is to be learned. It was also pointed out that job analyses of learning tasks must be done and measurement problems solved before valid comparisons between automated and conventional instruction can be made.

FOREWORD

Purpose

The concept of an automated teaching machine has been vigorously and enthusiastically investigated during the past few years. Educationalists, the armed forces, and business in general have shown a marked interest as to the implications that the automated tutor may have for their respective fields.

As it sometimes happens to a subject matter when the "state of the art" is in a period of fluid and rapid change, and when the investigators are literally on the frontiers of their science, issues become confused, problem areas are fuzzy and not sharply delineated, and zealous enthusiasm often is substituted for sound methodological evidence. In order to "clear the air", so to say, this review of theory and research was undertaken. During the past year several other reviews have appeared but while they have tended to be expository this current review is more critical in nature by singling out basic issues, evaluating problem areas, and recommending fruitful avenues of approach for further research.

Results

The results of this study can be summarized as follows:

1. Programs and the techniques of programming are crucial. The construction of items in order to ensure the elicitation of the correct response, the utilization of prompting, confirmation, and size of item steps evoke provocative questions with few answers readily available. In addition, such variables as the mode of response and the use of branching programs have yet to be definitively and systematically investigated.

2. There is yet to be found fundamental answers as to the properties of the teaching machine per se. Its display features, its capability of allowing self-pacing or machine pacing, and finally whether automated instruction will require supplementary forms of reinforcement to that already provided through confirmation -- these are questions that need empirical answers.

3. In comparing automated self-instruction with conventional instruction the general conclusion is that teaching machines promote learning. Due to

the paucity of well-controlled studies it is unwarranted to conclude that automated teaching is superior to conventional teaching at this time. Before a fair and adequate comparison can be made the further refinement of good programs and machines is necessary.

Implications

This review focuses in sharp detail problem areas. From a knowledge of such areas investigators will be able to answer basic issues, generate significant research studies, and, in general, empirically advance the "state of the art" Only thus can good programs and machines be developed.

Clayton K. Bishop

Clayton K. Bishop, Ph.D
Staff Psychologist

James J. Regan

James J. Regan, Ph.D
Head, Systems Psychology Division

TABLE OF CONTENTS

Section		Page
INTRODUCTION		1
I	PROGRAMS AND PROGRAMMING	3
	Techniques of Eliciting the Correct Response	4
	Prompting and Confirmation	8
	The Use of Explanations	10
	Size-of-Step	11
	Response Mode	13
	Branching	16
	Review Items	17
	Some Methods and Rules for Programming	18
	Summary	22
II	TEACHING MACHINES	24
	Repeating Incorrect Items	24
	Display Characteristics	24
	The Motivational Properties of Teaching Machines	26
	Pacing	27
	Feedback and Reinforcement	27
	Summary	28
III	AUTOMATED SELF-INSTRUCTION AND CONVENTIONAL INSTRUCTION	29
	Summary	33
IV	CURRENT TRENDS AND RECOMMENDATIONS	34
REFERENCES		37
TECHNICAL APPENDIX A		41

LIST OF TABLES

Table		Page
1.	Some Examples of Formal Prompts	6
2.	Some Examples of Thematic Prompts	7
3.	Items From Skinner's High School Physics Program. .	7
4.	An Illustration of Homme and Glaser's "Ruleg" System	20

INTRODUCTION

Psychologists and educators, in the laboratory and in the classroom, are becoming increasingly aware of the potentialities of automated teaching. This awareness involves, in some cases, an awakening to the possibilities of providing more efficient classroom instruction and of freeing the teacher from the routine chores of drill and exercise so long identified with teaching. In other cases, this awareness concerns the basic issues of teaching and learning, issues which include considerations of what to teach as well as how to teach.

The present fervor surrounding teaching machines apparently heralds the beginning of a new trend in educational research. Furthermore, there is reason to predict that changes in educational and training practices are likely to occur. While the trend is essentially a new one, automated teaching devices are not new. More than thirty years ago, Pressey (1926) devised a machine which was in many respects similar to the machines of today. This original machine was first conceived of as an automatic testing device, one which permitted the student to discover immediately whether his answers to test questions were right or wrong. Although designed for testing, the machine embodied one of the basic principles of learning, namely, the principle of reinforcement. According to Pressey, by immediately providing the student with information about the accuracy of his responses, the machines were making use of the "law of recency" and thus should be useful as instructional devices. Pressey even suggested that teaching machines could, in some instances, do a more efficient job of instructing than the "human machine." However, in spite of Pressey's enthusiasm and his vigorous attempts to demonstrate the effectiveness of automated teaching, it was not until 1954 or 1955 that a real movement toward automated instruction began.

The fact that the efforts of Pressey and his associates did not eventuate in any real educational trend was probably due to the absence of systematic programming techniques and to the emphasis placed on testing, rather than learning. There was no attempt to program these devices in terms of behavior control. Evidently such programming has had to await the growth of a science of behavior which could develop principles leading to the precise control of behavior. We would be rejecting history if we dismissed the early work of Pressey, for he did anticipate much of the work that is being done today. It would appear that Pressey was well ahead of his time, but neither science nor public opinion was able to encourage his efforts.

Today the situation is very different; the educational process is being scrutinized at all levels. The demand for formal education is far in excess of available resources. Technological advances in industry and in the armed forces have increased the demand for the training of highly skilled technicians. It is easy to understand why any innovation in educational and training practice might receive serious consideration if there were any hint that it might lead to an improvement in the efficiency of teaching.

Present-day educational needs have set the stage for the teaching machine movement, but of equal importance is the fact that the experimental psychology of learning has matured to the point where a technology of behavior control is more possible today than it was thirty years ago. This is not to

assert that the experimental psychologists are now confident that they can go from the relatively simple paradigms of the laboratory to the exceedingly complex variables of the classroom. However, many investigators would be willing to take the position that there is now a body of fairly precise quantitative laws describing the relationship among certain stimulus, response, reinforcement, and motivational variables. With these laws in mind, some investigators are willing to examine the educational process and attempt to develop a more efficient technology of education based on these laws. The leading exponent of the new technology of education is B. F. Skinner (1954, 1958), who argues persuasively for a technology based on the principles of a science of behavior, principles which, although derived from the laboratory, can be applied to the classroom.

According to Skinner, one of the most important of these principles is that of reinforcement. While reinforcement is a somewhat controversial concept in psychology, the principle can be defined empirically by simply describing the fact that certain environmental effects strengthen the behavior which has produced these effects. These environmental effects are typically referred to as reinforcements. Many reinforcements depend upon immediate physiological states such as hunger and thirst. For example, food will reinforce a hungry animal and water a thirsty animal. Many other reinforcing effects depend upon the animal's previous experiences. Novelty, praise, and material tokens represent part of a general classification of human reinforcements. The effectiveness of reinforcement depends upon the nature of the relationship between the behavior and the reinforcement. When the environmental effects which are to be reinforcing are fed back to the behaving organism quickly, reinforcement is effective. When there is a delay, sometimes of only a few seconds, the reinforcing effect is greatly weakened or even eliminated. Thus, in teaching a given response, the reinforcement must follow that response very closely. If the particular response is simply one of many responses occurring prior to reinforcement, much unwanted or irrelevant behavior will be reinforced. Therefore, in teaching complex behaviors, behaviors which include many different responses, it is necessary to use a series of progressive approximations, with each step in the series being reinforced so as to lead finally to the complex behavior in question. By giving reinforcement for each of the responses in the series making up the complex pattern, the desired behavior is gradually shaped. Furthermore, reinforcing each response assures that the learner will receive a large number of reinforcements, thus maintaining his interest and attention.

According to Skinner (1954, 1958), the most efficient way to take advantage of the principle of reinforcement and its ramifications in educational practice is to use the automated devices now generally referred to as teaching machines. Only in this way can one economically arrange the optimum conditions of reinforcement, the conditions of immediacy, precision, and frequency. With an automated device one can program the progressive approximations so important in the shaping of complex behaviors.

Skinner's position is probably the most influential single position in the teaching machine area today. However, his is not the only position; the field of automated teaching is fast filling with investigators from a variety of areas within psychology and education. This review will attempt to report on the present state of theory, research, and practice in the field of

teaching machines. The purpose of this review is to provide the reader with an overview of the state of present knowledge, as well as to acquaint him with the ongoing research. Many recently completed studies will be summarized, research in progress will be reported, and the various theoretical positions will be summarized.

At this time no single review of the teaching machine area can be exhaustive, since interest in this area is growing at a positively accelerating rate. Only a portion of this report is based on published papers and formal research reports. Many of the citations are unpublished papers, personal communications, and the replies to questionnaires sent out by the writer (Appendix A). The fact that many of the references are not formal reports, testifies to the embryonic nature of this research area, rather than to any lack in intellectual vigor.

The topics covered below have been classified into sections. Section I deals with issues and problems in the construction of teaching machine programs. Section II deals with problems relating to the design and construction of machines. Section III concerns control group studies comparing automated teaching to more conventional teaching methods. Section IV summarizes current trends and recommendations for future research.

I. PROGRAMS AND PROGRAMMING

A teaching machine is any device which can present systematically programmed materials while making efficient use of the principles of reinforcement. A teaching machine is not necessarily a hardware device; it can be a book, a set of cards, a mechanical or electrical unit, or an elaborate electronic apparatus. A device without a program is not a teaching machine. While there is some controversy over whether programming research is propaedeutic to machine-construction research, few investigators would quarrel with the position that the best machine is useless without a good program.

Within the general area of program construction there are many specific problems. Program writing is, according to the testimony of programmers, a most formidable task. It is formidable because it requires a discerning overview of the subject matter to be programmed, as well as a meticulous attention to details. Furthermore, there are few well tested techniques or rules for programming. According to Homme and Glaser (1959a), the programmer needs two systems: (1) a set of rules for analyzing the subject matter, and (2) a system for frame writing.[1] At the present time there are few, if any, systematic rules for analyzing the subject matter; neither are there well established systems for frame writing.

Some general rules about frame writing have been suggested by various writers (Skinner and Holland, 1958; Beck, 1959; Gilbert, 1958b; Smith, 1959;

---

1. A frame is an item in a set of many items. The term, frame, is used to describe an item because the typical teaching machine presents each item in a frame or aperture.

Homme and Glaser, 1959a; Carr, 1959). Some of these suggestions are based on experience in programming, some are frankly intuitive, and some very few rules are based on the results of experimental studies. These suggested rules will be summarized at the conclusion of this section. Before these rules are considered we need to review the kinds of problems which relate to programming. The following problem areas will be reviewed: (1) techniques of eliciting the correct response, (2) prompting and confirmation, (3) the use of explanations, (4) the size of item steps, (5) response modes, (6) branching, (7) the use of review items, and (8) some methods and rules for programming.

Techniques of Eliciting the Correct Response

In developing a teaching machine program, one of the first problems requiring consideration is that of eliciting the to-be-learned behavior. If the purpose of a program is to shape a series of responses into a complex act, then a means must be found to elicit each of the responses in the presence of certain discriminative stimuli.[1] While the possible effectiveness of prompts has generally been recognized, few formal studies have been specifically devoted to this problem. There are some comprehensive studies in progress whose purpose is to examine prompting methods.

A series of studies dealing with response prompting is under way at the American Institute for Research under the direction of Dr. A. A. Lumsdaine (1959a). The purpose of this project, sponsored by the Air Force Office of Scientific Research, is to investigate response-cuing factors in order to determine the optimal use of prompts. An analysis is to be made of the cuing functions furnished by contextual prompts, by prompting cues varying in physical clarity, and by prompts previously used. It is expected that this project will provide a taxonomy of prompting techniques, and that ultimately a psychometric scale of prompt strength which can be calibrated against a standard of physical variation in cue clarity, will be developed. Lumsdaine and his co-workers plan to perform several laboratory experiments to test the hypothesis that transition from maximal prompting early in learning to minimal prompting in the later stages will produce optimal learning. The technique of going from maximal to minimal prompting is essentially an extension of Skinner's (1958) principle of vanishing, in which cues are provided to insure the appearance of the correct response and gradually the cues are reduced ("vanished"). A good example of vanishing is Skinner's (1958) description of how to teach a student map labeling:

> When a student describes the geography of part of the world or the anatomy of part of the body, or names of plants and animals from specimens or pictures, verbal responses are controlled by nonverbal stimuli. In setting up such behavior the student is first asked to report features of a fully labeled map, picture or object, and the labels are then vanished. In teaching a map, for example, the machine asks the student to describe spatial relationships among cities,

---

1. A discriminative stimulus is one which sets the occasion for a response, e.g., the noon whistle sets the occasion for going to the dinner table; this behavior is reinforced by the presence of food.

> countries, rivers, and so on, as shown on a fully labeled map. He is then asked to do the same with a map in which the names are incomplete, or possibly lacking. Eventually he is asked to report the same relationships with no map at all. If the material has been well programmed, he can do so correctly.

Lumsdaine also plans to investigate special techniques of prompting which utilize variations in the physical properties of the prompting stimulus. Studies are to be done to test (1) the relative merits of having cue clarity adjusted by the machine rather than by the learner, and (2) the effectiveness of verbal and pictorial cues.

Skinner, Holland, and Meyer at Harvard University[1] are interested in the quantification of formal prompts. At present studies are under way at Harvard investigating the use of formal prompts in paired-associate learning. The technique is described as follows:[2]

> A German word and its equivalent is used in the paired-associate learning paradigm. The German word appears before the subject on an opaque screen. Under it is presented the English equivalent, badly defocused. The subject attempts to provide the English equivalent. When he is unable to do so he presses a switch which gradually causes the focus to improve for the English equivalent until he is able to provide the correct answer. As learning progresses, less and less focusing is required.

Based on their experience of two years work with teaching machine programs, Skinner and Holland (1958) have summarized the "Principal Features of a Useful Teaching Machine," giving special attention to the problem of eliciting the correct response. Two kinds of prompting techniques are discussed in this summary. One kind is referred to as a formal prompt and the other is called a thematic prompt. An explicit formal prompt is one which makes the to-be-learned response more probable, as in the case where the learner is given some of the key letters of the response word. Example I in Table 1 illustrates the key-letters prompt. This item was taken from Skinner's Program in High School Physics (1958) and has been modified to illustrate the formal prompt. Rhyming is another type of formal prompt shown in Table 1. The example given was taken from Skinner and Holland (1958). Holland (1959) frequently used prior definitions and examples as formal prompts in his Behavioral Science program. An example from that program is shown under III in Table 1. Holland also used textual material as formal prompts. For some sets of items, students were given some printed material to keep before them while they went through the set of frames. According to Holland, the distinction between giving the student this type of textbook material and simply having him read a text for memorization, is that the teaching machine program forces him to read the associated textual material for complete comprehension. Such comprehensive reading is assumed, not forced, with the typical textbook.

---

1. Information from questionnaire: Skinner, B. F. Harvard University, Cambridge, Mass., 1959
2. Quoted from questionnaire

Table 1

Some Examples of Formal Prompts

	Word to be supplied
I. Example of Prompting by Supplying Part of the Response Word.	
When we turn on a flashlight, an electric current flows through the fine wire in the b__b and causes it to grow hot.	bulb
II. Example of Prompting by Rhyming	
"Nine times seven and just one more Is eight times eight or _____ "	64
III. Example of Prompting by Prior Definitions	
"A technical word for reward is reinforcer. To reward an organism with food is to _____ it with food."	reinforce

A second type of prompt, different from the contextual, formal prompt is the thematic prompt. Items which make use of the student's prior knowledge, his word habits, his indicating categories and syntax, are thematic prompts. One kind of thematic prompt is the "lead in" item which utilizes common knowledge or phrases with high association value for the to-be-learned word. This is illustrated in Example I, Table 2, with an item from Holland's Behavioral Science Course. According to Holland, the correct response, "reflex," was given by 96% of his students, since it is fairly common knowledge that the knee jerk is a reflex.

Another type of thematic prompt makes use of categories like "opposites" or analogies. Example II in Table 2 contains an analogy item taken from Lumsdaine (1959b) and an "opposites" item taken from Holland's Behavioral Science Course. Grammatical construction can serve as a thematic prompt. Whether a response is to be a verb, noun, adjective, or adverb can often be determined by the structure of the sentence and the place of the blank; or by the nature of the article, a or an; or by supplying word suffixes such as -by or -ed.

Thematic prompts can be acquired in the course of learning. By introducing a new word in a series of frames and giving the word a significant place in each of the frames in the series, the probability of the student's responding with that word in the last frame of the series is increased. An example of this technique is shown in Table 3, where items 13, 14, and 15 from Skinner's (1958) High School Physics Program are reproduced.

Table 2

Some Examples of Thematic Prompts

	Word to be supplied
I. Example of Thematic Prompting Which Uses Common Knowledge.	
"A doctor taps your knee (patellar tendon) with a rubber hammer to test your _____."	reflex
II. Examples of Thematic Prompting Making Use of Analogies and "Opposites"	
"Just as the rise of temperature increases gas pressure, a rise in filament temperature will _____ electron emission from the filament of a vacuum tube."	increase
"Reinforcement which consists of presenting stimuli (e.g., food or water) is called positive reinforcement; reinforcement which consists of terminating stimuli (e.g., loud noises or painful stimuli) is called _____ reinforcement."	negative

Table 3

Items From Skinner's High School Physics Program

	Word to be supplied
13. "An object which emits light because it is hot is called 'incandescent.' A flashlight bulb is an incandescent source of _____."	light
14. "A neon tube emits light but remains cool. It is, therefore, not an incandescent _____ of light.	source
15. "A candle flame is hot. It is a(n) _____ source of light."	incandescent

The items in Table 3 illustrate how the word "incandescent" is introduced in one frame, repeated in a second frame, and called for as the response to the third frame.

Prompting and Confirmation

Closely allied to the problem of eliciting the to-be-learned response is the problem of strengthening the response, that is, increasing the probability of the occurrence of the response in the presence of the discriminative stimulus. According to the principles of reinforcement, this strengthening is done by following the correct response with an immediate reinforcement. In the teaching machine situation the reinforcement is, most often, immediate confirmation that the response in question is correct. According to some investigators (Skinner, 1958; Porter, 1958; Lumsdaine, 1959b), immediate confirmation is one of the major features of good teaching machine programs. However, some recent studies have shown that prompting without apparent confirmation is an effective technique. Irion and Briggs (1957), using the Subject Matter Trainer (described in detail by Briggs, 1958), found that a coaching or prompting technique was most effective in serial learning, paired-associate learning, and problem solving. With this technique, called the "quiz mode" or "coaching mode," the subject responds to an item by pressing a special button on the device which immediately produces a light next to one of the correct answers in a group of 20 possible answers. Thus, the subject is told the correct answer without having to give his own response.

The second most effective technique in the Irion and Briggs study was the "modified-quiz mode" or "single-error permitted mode" in which the correct answer is signaled after the first try by the subject. Here the subject is given the correct answer after he has made one response. The "practice mode" and the "single try mode" were the least effective techniques. In the "practice mode" the subject keeps responding until the correct response is made. In the "single try mode" the subject responds only once to each problem. He is signaled when he is correct and when he is incorrect, but when incorrect he is not informed about the correct response.

Cook and Kendler (1956) and Cook (1958) found that response prompting produced faster paired-associate learning than did response confirmation. The prompting technique consisted of presenting the response member of the pair before the subject had made his response. Kopstein and Roshal (1955) also found more rapid paired-associate learning with a response prompting technique, in which the stimulus and response members are presented simultaneously, as compared to a confirmation technique, in which the stimulus members precede the response member in time.

Bryan and Schuster (1959) found a guidance condition to be superior to a no guidance condition in teaching electronic trouble shooting. Subjects who were given a list of trouble shooting steps to follow, that is, who were prompted as to what responses to make, achieved higher scores on a criterion test than did subjects who had to find the correct response without guidance.

The above studies, dealing with prompting and confirmation, present a fairly consistent picture favoring the use of prompts. However, these studies contain confounding features which make some of the conclusions equivocal. The Cook (1958) study and the Irion and Briggs (1957) study were both confounded by uncontrolled variables. Sidowski, Kopstein, and Shillestead (1959) point out some possible flaws in Cook's procedure which they are attempt-

ing to correct in their own studies of prompting and confirmation. One such flaw was that Cook presented a test trial after every third training trial, thus interspersing, at frequent intervals, a type of confirmation trial. According to Sidowski et al., the combination of prompting and confirmation may have accounted for Cook's findings. Another variable not well controlled in the Cook study was the delay between the stimulus and the response presentation under the anticipation or confirmation condition. Sidowski et al. suggest that a procedure which would present the stimulus member as well as the response member after "practice" may offset the superiority of prompting. Saltzman[1] reports that after the subject has been shown the stimulus member of a pair and has made his response, it is more efficient to present both the stimulus and response word together as confirmation than it is to present the response member alone.

In the Irion and Briggs study a number of factors could have contributed to the superiority of the coaching mode. The authors themselves indicate four such factors: (1) because no time is lost in searching for the correct response, more trials per unit time are accomplished; (2) overt wrong responses are avoided; (3) aversive signals for wrong responses are eliminated; and/or (4) there is a close temporal contiguity between the stimulus, response, and the correct signal. In order to determine the effects of coaching or prompting, these four factors need to be controlled. In terms of the manner in which the study was performed, the superiority of the coaching mode could have been the result of any or all of the above factors. A confirmation technique which permits more practice, fewer errors, eliminates aversive signals, and insures contiguity between stimulus, response, and confirmation, might well prove superior to a prompting technique. In fact, it can be argued that the small step, gradual shaping programs advocated by Skinner are constructed with these factors in mind. If comparisons are to be made between prompting and confirmation techniques, these comparisons should only be made using carefully finished programs.

The prompting or coaching mode is not necessarily opposed to the confirmation technique. It is easy to envisage a combination of prompting and confirmation and in many instances Skinner and his co-workers use such combinations in their programs. Gilbert (1958a) has suggested a method of responding which includes construction, confirmation, recognition, and prompting. The learner first constructs his response and then he selects from a number of alternatives the one which most closely resembles his response. The machine informs the learner whether his choice was correct.

It is likely that the type of response sequence described by Gilbert occurs covertly with many of the multiple-choice programs. If we assume that the subjects in the Cook, Irion and Briggs, Kopstein and Roshal, and Bryan and Schuster studies, referred to above, may have constructed a response prior to attending to the prompt, then prompting becomes a form of confirmation. If this is the case, then the distinction between prompting and confirmation depends upon the relation between the learner's overt response and the prompting or confirming cue. When the cue occurs before the learner's overt response, it is a prompt; when it occurs after the overt response, it is a confirmation.

---

1. Personal communication, 1959

Since we are seldom able to know the learner's covert responses, the distinction between prompts and confirmation can only be made in terms of the learner's overt behavior.

With regard to the time relationship between response and confirmation in learning, there is generally little disagreement that immediate confirmation is desirable, but there has been little research on this problem using teaching machines. Meyer[1] reports that immediate confirmation made a considerable difference in the ability of junior high students to handle a vocabulary building program. Subjects who were trained with immediate confirmation made higher scores on a criterion test. However, Porter (1957) points out that immediate confirmation is often confounded with frequency of confirmation, and the latter variable may be accounting for the effects of immediate confirmation. It should be a simple matter to separate the two variables, frequency and immediacy, in order to determine the relative effects of either. Moreover, there are a number of different applications of immediate and delayed confirmation which should be tested. For example, what advantages are there to confirming individual responses as compared to confirming a group of responses? Perhaps the shaping of response sequences can be best accomplished by going from immediate confirmation for individual responses to delayed confirmation, or confirmation for groups or sequences of responses. Anderson[2] reports that she plans to investigate the effects of knowledge of results given after each response and given after several responses.

The Use of Explanations

Some teaching machine programs simply confirm whether or not the learner is correct by presenting the correct response, while other programs use the confirming mechanism to provide explanations and further information. Crowder (1959) uses explanations in his scrambled book technique. When a learner chooses the correct answer, he is told that he is correct and is then usually given some brief explanation why that particular choice is correct. If he is wrong, he is told, given an explanation why he is wrong, and then told to go back to the item and to try again. For Crowder, explanations are important, since he considers the communication process "to be the essential process in the teaching of human beings" (1959, p. 114). When an error has occurred, the communication process is revised using explanations; when the learner is correct, explanations are used to extend the communication process.

Gilbert (1958a) has also suggested that confirmation might be used as a source of information, but he does not emphasize the role of explanatory material as does Crowder.

A study by Bryan and Rigney (1956) presents some information about the use of explanations in a training task. These authors found that subjects who were provided with explanations to accompany each alternative in a "pull tab," multiple-choice device, were superior on a multiple-choice criterion

---

1. Questionnaire: Meyer, Susan R. Harvard University, Cambridge, Mass., 1959
2. Questionnaire: Anderson, Nancy S. University of Maryland, College Park, Md., 1959

test when compared to control subjects who received no explanations. Another study by Bryan and Schuster (1959), referred to earlier, also demonstrated that groups receiving explanations during training on an electronic trouble shooting task, performed better on a criterion task than did groups receiving no explanations.

Size-of-Step

In the design of teaching machine programs much emphasis has been placed on producing short, easy steps so that the learner can gradually, without making errors, reach the desired level of learning. This is the technique of shaping behavior by successive approximations that is espoused by Skinner (1954, 1958). Skinner maintains that items should be worded and sequenced so as to permit the learner to proceed without making errors while receiving a high percentage of reinforcement. However, there is some controversy regarding the desirability of using very small steps in teaching machine programs.

Among those investigators taking issue with exclusive use of short, easy steps is S. L. Pressey (1959), who questions the validity of the small-step program. Pressey suggests that having to experience several trials before being successful might be educative and that exposure to incorrect responses might make for more efficient learning. Pressey is also concerned about the use of small steps with superior students. He implies that detailed step programs might impede the progress of the bright student by reducing his motivation. Smith (1959) also has suggested that there may be some advantage to inducing errors. He points out:

> Apparently, in an imperfect program, motivation runs high, so much so that even "honest" children try to cheat. It seems as though the possibility of failure enhances the reinforcing effect of a successful response. Thus failure on some items may contribute to the effectiveness of learning.

Skinner (1958) has anticipated arguments like those of Pressey and of Smith. He suggested that inducing errors is a technique concerned with motivation and not with learning. The use of difficult items which force errors is based on the principles of aversive control, rather than on the principle of control by positive reinforcement. According to Skinner, easy items will simply permit the quick student to proceed more rapidly; there is little reason to prejudge such items as being boring or nonmotivating.

Interest in the size-of-step problem has resulted in some experimental studies. Evans, Glaser, and Homme (1959) studied size-of-step using part of a program designed to teach conversion to number bases other than ten. Their teaching machines were index cards and a type of programmed textbook. Four sequences of the program were used in the experiment: (1) 30 steps, (2) 40 steps, (3) 51 steps, and (4) 68 steps. The fewer the number of steps, the larger was the size of each item step. Care was taken to see to it that the essential points in the development of the subject matter were included in each sequence. Four independent groups of five graduate education students each were given the sequence. The criterion measures of learning were immediate and delayed written tests. The results indicated a relationship between number of steps and number of errors on the tests; the smaller the number of

steps, the fewer the number of test errors. The smaller steps were also associated with less time per step and fewer errors in the course of learning. However, the group given the 68 step sequence was slightly inferior to the group taking the 51 step sequence, suggesting that there may be an upper limit beyond which it is no longer efficient to extend the number of steps.

Coulson and Silberman (1959) performed a multivariable study with one of the variables being step size; the other variables were response mode and item sequence. The program consisted of the first 104 items from Holland's Behavioral Science Course. For the small step-size groups 104 items were used and for the large step-size groups 56 items covering the same material included in the 104 items were used. Human experimenters served as simulated teaching machines. The experimenters were seated behind a screen where they could control item presentation and confirm the subjects' responses. One-half the subjects in each step-size group constructed their responses by writing them on a slip of paper. The remaining half of the subjects selected their answers from a set of five choices. The step-size, response mode groups were further divided into those subjects who branched and those who did not. For the subjects who branched, certain items were eliminated if particular key items were answered correctly on the first try. In the nonbranching condition the subject went through all the items in order. Thus, in the Coulson and Silberman study there were a total of eight conditions, with ten subjects being run independently in each of the conditions. The subjects were students from an elementary psychology course. A control group of 104 subjects was simply continued in the course and was not given any of the teaching machine programs.

The criterion tests consisted of 19 constructed response questions and 17 multiple choice questions. The test was administered immediately after learning and again three weeks later. In addition to the tests, the time required to complete training was included as another performance measure. The results of the Coulson and Silberman study showed:

(1) The experimental group was superior to the control group on both of the immediate tests. The authors point out that this does not indicate that teaching machine training is superior to conventional classroom instruction because the control group was not exposed to the same concepts as the experimental group. According to the authors, the control group simply served as a type of no training group.

(2) Small item steps required more training time, but also yielded significantly higher test scores than did large item steps on the constructed response test.[1]

(3) The multiple-choice response mode took less training time than the constructed response mode, but no significant differences in test scores were found.

---

1. These results refer only to the immediate criterion test. The later criterion test was used as a measure of retention. These test scores revealed a significant improvement (practice effect) between the first and second administration of the multiple-choice test, but no significant improvement for the constructed-response test.

(4) The branching conditions required less training time than non-branching, but there were no significant differences on the criterion tests. A significant interaction was found between response mode and branching on the constructed response test. This interaction was due to the high mean score obtained by the constructed-response, nonbranching group.

While the results of the two studies reported above could be interpreted as lending support to Skinner's position regarding size-of-steps, the data are not sufficiently conclusive to halt further debate about the issue of step-size. Evans, Glaser, and Homme (1959) used only five subjects in a group and Coulson and Silberman (1959) used only ten subjects in a group. In neither experiment was there a sufficiently large or heterogeneous sample to shed light on the question of step-size and individual differences raised by Pressey (1959).

Both studies fail to give serious attention to the fact that the small-step groups spend more time with the material to be learned. In the Coulson and Silberman study the small-step group spent an average of 57.3 minutes with the material, while the large-step group spent an average of 41.5 minutes. It is possible that the 16 minute difference between the two groups could account for the difference in test performance. Time differences of this order were not observed for any of the other conditions. If step-size is an important variable in its own right, it deserves more precise experimental treatment than that given by either the Evans, Glaser, and Homme study or the Silberman and Coulson study.

Closely linked to the question of size-of-step is the question of error frequency. Evans et al. found, as was expected, that with small steps fewer errors were made in the course of learning. Although they did not report the frequency of errors, it is probable that similar relationships occurred in the Coulson and Silberman study. It is not surprising to find that easy, small-step items are associated with relatively few errors. However, closely linked as errors and size-of-step are, they are not rigidly bound. It should be possible to separate the two variables by using small-step programs seeded with error-producing items. Pressey (1959) has suggested that errors may be useful; Skinner (1958) claims they are not relevant to the learning process as such. It may be informative to test the effects of errors per se, particularly if we wish to distinguish among the successive approximations (small steps) technique of teaching, the possibly aversive (motivational) effects of errors and the information giving aspects of errors.

Response Mode

An issue of considerable interest in teaching machine programming is the question of response mode. Skinner (1958) maintains that the response should be constructed rather than selected from a group of alternative choices. He argues that the purpose of teaching is to develop recall rather than recognition. The teacher wants the student to make a response. Skinner further maintains that the presence of plausible wrong responses among the available choices can impair the gradual shaping process by strengthening undesired responses. On the other hand, Pressey (1959) asserts that the multiple-choice mode is an efficient learning procedure.

The Coulson and Silberman study described above compared the constructed-response mode with the multiple-choice mode. No significant differences were found between the two response modes, although the multiple-choice mode took less time to complete than did the constructed-response mode.

Fry (1960) found that subjects who were trained to make constructed responses were superior to those who were trained to make multiple-choice responses. He taught ninth grade students Spanish words and phrases under the constructed response and under the multiple-choice conditions. Significant differences favoring the constructed responses were found using constructed response test items. Multiple-choice test items failed to reveal any differences between the two learning modes.

The problem of response mode is not confined to the issue of constructed versus multiple-choice responses. Questions have been asked about the need for making an overt response and about the type of overt response, e.g., vocalizing, writing, pressing a key, and so forth. Evans et al. (1959) performed a study using part of a sequence on "Fundamentals of Music." Two groups of undergraduate psychology students were used, with the subjects in one group instructed to respond to each item by writing their answers, and the second group was told not to make any overt response. The results of this study indicated that the nonwriting group took less time to complete the sequence and did slightly better on a criterion test, although not significantly, than did the writing group.

Silverman (1959a) found no differences among four different response mode conditions. The conditions were: (1) writing the response at each step, (2) writing and vocalizing the response at each step, (3) vocalizing the response at each step, and (4) neither writing nor vocalizing the response. The subjects in the fourth group were instructed to "think." All the subjects were paced, that is, they were exposed to each item for a similar period of time in order to keep item exposure constant for each of the groups. The subjects were college undergraduates; the program material was part of a course in "Basic Electricity"; and the criterion measure was a performance test given immediately after learning.

Zuckerman (1959) is engaged in a project at Brooklyn College which entails a study of response mode. Comparisons are to be made between constructed responses, multiple-choice responses, and true-false responses, and between a sequenced presentation and a random presentation. While it is difficult to imagine a rationale for random presentation of items, the interaction effects involving response mode and type of presentation might shed some light on the logic of programming.

The question of response mode leads to the further question about the behavior to be learned. When Skinner asserts that recall is preferred over recognition, he is speaking about a general performance class. However, what is it that the learner should recall? Is the phrase, word, or symbol which the subject constructs the important response, or is it the information contained in the item (the verbal sequences) that should be retained? Although these questions have not been given the attention they deserve, they have not been entirely ignored. Holland (1959) has discussed the problem of controlling the learner's observing and echoic behavior. He points out that

often when the programmer omits a word, phrase, or symbol which the learner is to construct, the programmer's purpose is simply to insure a careful reading of the item. Holland states, "Only those parts of an item which must be read to correctly complete a blank can safely be assumed to be learned." Apparently, Holland's position is that we are not necessarily trying to teach the constructed response; but by having the learner make such a response, we are insuring his attention to the verbal sequences and we can assume he is engaging in "covert echoic behavior." Thus, according to Holland, the control of attention or "observing behavior" is especially important. The "covert echoic behavior" is also important, but the implication is that this behavior simply follows the behavior of attending to the material. By using blanks to be completed, attention is controlled. Furthermore, by having the learner make discrete responses, reinforcement can be given in the form of confirmation.

Holland's position suggests that the blanks, the to-be-constructed or chosen responses, are themselves not the important feature of teaching machine items. If a learner is highly motivated, he will probably read the step-wise sequence carefully, paying close attention to the details, and performing as effectively as a learner who was given the same material with blanks to be completed. While there are no data which directly support the view that the blanks within items simply insure attention, there are some findings which point out that such blanks may not always be necessary. Kormondy (1959) found that a sequenced textbook with no blanks is as effective as a sequenced workbook containing blanks. Kormondy's study is described in more detail in a later section (p. 32). It is sufficient to point out here that the study was performed with college students who knew that they were participating in an experiment.

Silverman (1959a) found that subjects who simply read items presented sequentially by means of a teaching machine (Skinner disk machine) showed better learning than subjects who took the usual fill-in-the-blank program. The subject matter (basic electricity) was identical for both groups of subjects; the only difference in the two group conditions was the requirement for one group that they construct answers to the blanks.

A possible part explanation of the Kormondy and Silverman findings is that in both experiments a "Hawthorne" type of effect (Kitson, 1922) may have been operating. In both studies the subjects knew they were in an experiment; the Silverman subjects worked in a laboratory setting and the Kormondy subjects remained in the classroom. It is possible the knowledge that they were experimental subjects was sufficient to motivate them to make the careful observing responses so important to learning. Furthermore, the motivation provided by the experiment may have been strong enough to keep them working without the motivation provided by specific reinforcement for each response. The fact that Silverman's "no-response" (no fill-in items) showed better learning may have been a function of the program. Perhaps the fill-in responses were distracting; a subject who simply read the items making echoic and rehearsal responses may have acquired a greater range of information than one who diligently attended to the responses.

Another possible explanation of Kormondy's and Silverman's findings is that the key feature in teaching machine programs may actually be the

small-steps and the sequenced presentation. If the learner is sufficiently motivated, it may be unnecessary for him to make the specific fill-in responses which insure attention, or it may even be detrimental to make these fill-in responses.

There is much to be learned about the kinds of responses that programmers can and should teach. While some investigators have achieved a measure of success in programming with observing behavior in mind, there still remains the question of item content. This question has important implications when we consider the purpose of teaching. Do we want the learner to be able to recall a list of facts, solve problems, transfer his training to new problems, and/or originate new questions? Holland (1959) indicates that he has not ignored the questions about what to teach. He states, "However, little academic education is simple discrimination. More often, it is abstraction or concept formation." He goes on to point out a method of teaching abstraction by programming items which contain many examples, each with a common property but also containing a wide range of other properties. By gradually progressing to more and more complex examples, abstraction is established.

Branching

The majority of teaching machine programs presents the subject material in a fixed sequence of steps; there is no provision for varying the sequence to meet the needs of an individual student. According to Skinner, a carefully programmed sequence does permit flexibility in terms of the speed or rate of learning. The quick learner simply proceeds through the program more rapidly than does the slow learner. For Skinner there is little to be gained by dropping or adding items for particular students.

Crowder (1959) has suggested that flexible or "intrinsic programming" might be advantageous in that each program is adapted directly to the individual student's present state of knowledge. An adaptable program provides branches which depend upon the student's response to a particular set of items. Obviously a branching program is more difficult to present than a standard program. If a hardware device is used, the technological problems are many. But if a textbook device is used, the presentation is relatively simple. Crowder's "scrambled book" (1959) is an example of an intrinsically programmed device in which the student proceeds on the basis of his response to each item. When his response is correct, he is directed to the next item in that series. When his response is incorrect, he is directed to an explanatory item and is then instructed to return to the previously failed item to try again. Crowder's approach is clearly different from Skinner's, in that it depends less on the successive approximations or shaping technique and more on the automation of the human tutoring process. Crowder's programs attempt to imitate the tutor, while Skinner's programs are based on a new technology of education derived from laboratory principles of learning.

The study by Coulson and Silberman (1959), described above, investigated the branching variable. Under the branching conditions the program sequence was divided into subsets. If the subject answered a certain key item correctly, the remaining items in that subset were eliminated. Thus, in the branching condition the exact number and sequence of items were not fixed but depended upon the subject's performance. In the nonbranching conditions

the subject simply proceeded through the items in order. Coulson and Silberman found no significant differences in performance between the branching and nonbranching conditions. However, it must be pointed out again that the time differences may have confounded the results. The subjects learning under the branching condition spent less total time with the material than did the nonbranching subjects; this may have worked to their disadvantage on the performance test. A better test of the effects of branching would equate total exposure time for the branching and nonbranching groups. In such a test, the subjects would be kept at the learning task for the same length of time; the branching subjects would probably cover more material, but all subjects would be tested on an equivalent section of the material. If, under the branching condition, subjects learned more and retained as much or learned as much and retained better than under the nonbranching conditions, then there would be grounds for concluding that branching was the more efficient programming technique for that particular subject matter. However, such a conclusion would be valid only if the nonbranching program was a well constructed, small-step program. Only in this way could a meaningful comparison be made between the branching and nonbranching methods.

Review Items

One of the basic questions of program construction that is only now beginning to attract attention, is the question of review items. How much review and repetition of material is necessary for long-term performance? Many of the early programs, particularly those employed by Skinner and his associates, made use of an item drop-out feature, in which correct items were eliminated and only the incorrect items were repeated. In these programs there was little or no provision for the review of the correct items. The rationale for such a procedure is that in a well constructed program which makes use of the principle of gradual approximation, the student is so well prepared for the next item in the sequence that there is little possibility of his making a wrong response. If the program proceeds in a cumulative fashion, then each step in the program assumes that what has gone before is learned. When the question of long-term retention is raised, the corollary question of the place of review items is introduced.

Kormondy (1959) has studied some aspects of the review problem with a program in "Introductory Genetics." Using booklets rather than hardware devices, he compared five groups of approximately 40 subjects each. One group took a Skinner-type program of 497 items. A second group took the identical program, except after the fourth, seventh, eleventh, and thirteenth lessons this group took a review exercise of approximately 25 items. Including the review items, the second group's program consisted of 588 items. The review items were, in some cases, syntactical rearrangements of earlier items and, in other cases, reworded items or new contextual situations. Kormondy's preliminary findings indicate no differences between the review and no-review conditions.

Another study now in progress is concerned with the question of review items. Ferster (1959), at the Indiana University Medical Center, is studying the role of review material in continuous programming by comparing two groups: "The first group will have the first 22 lessons without repetition, 23-34 with repetition, 35-55 without repetition; the conditions of the

second group will be the reverse of the first." Group differences are to be compared at the end of the 22nd, 34th, and 55th lessons. Ferster's program consists of items from the Harvard Behavior Science course.

Some Methods and Rules for Programming

Earlier in this section we mentioned that a number of suggestions for program construction have been offered by programmers. Some of these suggestions are summarized below. Few, if any, are based on experimental findings; if experiments have been involved, they are described above. Many of the suggestions are based on experience, a somewhat ephemeral variable, and on intuition derived from experience. Whatever may be the intent of the various writers, we shall not regard their suggestions as pronouncements, but shall, instead, examine them with heuristic purpose in mind.

Skinner and Holland's summary of the "Principal Features of a Useful Teaching Machine" (1958) contains some suggestions about programming. Some tentative recommendations are made regarding the length of item sets as well as the length of individual items. According to the summary, the time a student can work with a machine program may have no relationship to the normal length of a classroom session. It is suggested that a set of about 30 frames may be about the right length for the kind of material used in the Harvard Behavior Science course. A good student might complete several sets in a single session, while a slow student will have sufficient time in a one-hour session to complete one set.

The suggestion is made that where an item is lengthy care should be taken to insure that the text of the item is completely relevant to the right answer.[1] The question of the number of responses called for within a given frame is also considered, since two or more blanks can be used in the same item. Where multiple blanks are used, the item writer is cautioned to maintain clear syntax. It may be desirable, according to the Harvard group, to eliminate some of the syntax when the subject has been well prepared, and by so doing put the response more under the control of the subject's own stimuli.

Skinner and Holland recommend that in composing programs:

(1) The programmer must not appeal to material not included in earlier parts of the program. At each step the programmer should inquire whether the student is ready for the next step. Of course, the student is assumed to have some verbal ability before he starts a program which entails verbal behavior. Each frame should be intelligible by itself, but a <u>set</u> of frames assumes a common discourse.

(2) The programmer should list the terms to be covered in each set of frames, the principles involved, and a wide range of examples illustrating the principles. This suggestion is similar to that made by Homme and Glaser (1959a).

---

1. This suggestion is not completely consistent with Holland's (1959) discussion of blanks as maintainers of attention.

Beck (1959) has found it useful in preparing programs to consider the various ways of controlling responses which he calls "teaching rules." "Teaching rules" can differ in the way they control behavior and the degree to which they restrict behavior. Ways of controlling behavior include the use of stimuli other than discriminative stimuli. Some examples are stimuli for copying, the use of analogies, and deductive rules. Copying rules are useful in teaching specific responses. However, Beck indicates that often it is not a specific response that is desired but a large number of conditional responses.[1] He further points out that a deductive teaching rule would be most efficient where one is interested in either teaching deductive thinking itself or in teaching a class of conditional responses. In problem solving, where one wishes to teach a student to make analogous responses in related situations, a teaching rule based on analogies would be used.

In order to facilitate the development of principles of programming, Gilbert (1958b) has proposed a functional classification of item types. This classification system contains eleven categories as follows:

(1) Lead-in Items are items that neither give new information nor require the rehearsal of old skills. Their function is to orient the student to a problem and prepare him for new information.

(2) Augmenting Items are items that supply new information to a student but do not require him to make any relevant response. Responses required in such items may usually be of a kind that will help insure the students' reading the item.

(3) Interlocking Items are items that require a student to review established skills while new information is being presented.

(4) Rote Review Items present a problem identical to one earlier presented.

(5) Restated Review Items require the rehearsal of a skill where the problem is restated in a different syntactical arrangement.

(6) Delayed Review Items allow for further practice distributed in time. They will be different from other items only in time of presentation.

(7) Fading Items are items that not only require the student to review what has been presented to him, but also have information withdrawn from them from item to item.

(8) Generalizing Items present a verbal statement pointing to a common characteristic of several specific problems already presented to the student.

---

1. The term "conditional responses" (as the term is used by Beck) refers to the fact that one response is made under one kind of condition and a different response is made under different conditions.

(9) Specifying Items are items which exemplify a general rule in specific cases.

(10) Dovetailing Items require the student to make separate responses to separable stimuli that otherwise become confused.

(11) Subject-Matter Classification: Each item should be further classified with respect to content.

In classifying his own algebra program, Gilbert found that the majority of the items was restated review items. There were no rote review items nor delayed review items. According to Gilbert, rote review items probably should be avoided (for the sake of syntactical variation), except in the case where memorization of a short verbal statement is desired.

Smith (1959) has speculated about the characteristics of successful programs and programmers. With regard to successful programmers, he makes the somewhat unique suggestion that the best programmers are introverts and that this should be taken into account when selecting potential programmers. With regard to programming, Smith makes many suggestions, most of which must be read in the context of his paper where he cites illustrations.

Homme and Glaser (1959b) have proposed a set of systematic rules for programming. They suggest that most teaching involves the presentation of principles or abstractions (rules) and instances or special cases of these abstractions (examples). They call their system the "ruleg" system because it makes use of rules and examples. Rules are used to get the behavior in question to be emitted in the presence of the discriminative stimulus. The rule may serve as a prompt for a response such as completing an example. Incomplete examples can be used to extend specific discriminative control, and incomplete rules can be used to extend a more general discriminative control. As an illustration of the "ruleg" system, Homme and Glaser have presented a modification of part of Skinner's High School Physics Program (1958). The first two items of their illustration are shown in Table 4. Item 1 in

Table 4

An Illustration of Homme and Glaser's "Ruleg" System

	Word to be supplied
1. To "emit" light means to "send out" light. For example, the sun, a fluorescent tube, and a bonfire have in common that they all send out or _____ light.	emit
2. A firefly and an electric light bulb are alike in that they both send out or _____ light.	emit

Table 4 consists of a rule and an incomplete example; item 2 consists of an incomplete example.

Homme and Glaser claim that the "ruleg" system has been most helpful to them. It eases the programming process; it results in some economy in programming; it facilitates talking about programs; and it facilitates instruction in programming. The authors report that they have been able to instruct undergraduate students in programming "with encouraging results."

In reviewing some of the current concepts in the area of self-instructional devices, Carr (1959) devotes much attention to the process of programming. He isolates what he considers to be the relevant variables in programming by distinguishing among terminal, transitional, and initial stimulus-response connections. The process of teaching, according to Carr, consists of moving from the initial, through the transitional, to the terminal connections. The initial connections involve the eliciting stimuli and the responses which approximate the to-be-learned responses. The transitional connections are the steps by which the initial connections are changed. The initial response is replaced by a series of transitional responses and the initial stimulus is gradually eliminated, until finally the terminal stimulus-response connections are firmly established.

Making use of the concepts of initial, transitional, and terminal stimulus-response connections, Carr describes five possible sources of error in programming: (1) The programmer may incorrectly specify the terminal stimulus-response connections by not obtaining a clear picture of the subject matter to be taught. (2) The programmer may not estimate the extent of the initial stimulus-response connections correctly. An example of this would be over- or underestimating the learner's level of knowledge prior to training. (3) The programmer may not allow for sufficient strengthening of the transitional or terminal connections. This would occur in programs where sufficient care had not been taken to insure sufficient training on any one aspect of the subject material. (4) The progression of the transitional connections may be awkward or defective, such that complex responses are introduced before elementary skills are mastered. (5) The programmer may make premature progressions through the transitional connections, thus moving too rapidly for the learner.

Carr points up five classes of independent variables which he feels may influence the effectiveness of self-instructional programs. These variables are relevance, availability, sequence, stepping, and maintenance. Relevance refers to the specification of the to-be-learned task. Availability refers to the learner's initial repertory of behavior: the responses which he can make prior to training. Sequence concerns optimum order of presentation. Problems relating to straight line or branching type programs are included in the sequence variable. Stepping concerns the rate of progression from one transitional connection to the next. Maintenance refers to the amount of conditioning required of both transitional and terminal connections to insure adequate learning of these connections.

In spite of the dearth of experimentally tested rules, there is agreement among programmers. For example, Homme and Glaser's "ruleg" system is related to the principles espoused by Skinner and Holland and by Beck. Gilbert's classification and Skinner and Holland's suggested principles are

used by Carr in his initial, transitional, and terminal connection schema. This agreement among investigators is, of course, not fortuitous. Each has been influenced by the other and they have not arrived at their various schemas independently. If they had, we could be more confident in their recommendations. Only experimental tests can ultimately determine the validity of these various recommendations. These tests must entail comparisons among programs, for the only criterion by which to judge programming rules is the effectiveness of the program.

Summary

Conclusions.- A great part of the interest in the teaching machine area is focused on programs and programming. Without good programs, the most elaborate device is useless. But good programs have not been easy to develop because there is a dearth of systematic information about programming. However, there is no shortage of provocative questions about programming methods. Many of these questions are raised in this section; few are answered.

One of the basic problems in programming is that of constructing items that insure the eliciting of the correct response. In spite of the apparent fundamental nature of the problem, there has been much description of possible response-prompting techniques and few experimental studies. However, there have been some studies dealing with the difference between confirmation and prompting. These studies seem to show that prompting is more effective than confirmation, but they contain many shortcomings. Consequently, any conclusions about the comparative effectiveness of prompting must await the results of more carefully controlled studies.

Some investigators consider the use of explanations in teaching machine programs to be important, but only two studies (Bryan and Rigney, 1956; Bryan and Schuster, 1959) have demonstrated the value of explanations.

The size of the item step has attracted considerable attention among programmers. Skinner advocates the small step; Pressey suggests that the large step has a distinct place in programming, particularly if we consider individual differences. The two studies (Evans, Glaser and Homme, 1959; Coulson and Silberman, 1959) which have investigated size-of-step show that small-step programs produce better performance on criterion tests. However, the two studies failed to sample a large enough group of learners to shed light on Pressey's argument. Furthermore, the two studies confound error frequency and size-of-step. These two variables require separate study.

There is little agreement regarding the different response modes used in teaching machine programs. The findings to date do not justify any conclusions that multiple-choice responses are less effective than constructed responses, or that writing the response is better than speaking the response, or better than not responding overtly. With regard to response mode, the question is raised about the to-be-learned response: What is the response we are trying to teach? This question requires considerable clarification, particularly in view of the studies which fail to demonstrate that the fill-in responses advocated by Skinner are necessary for efficient learning.

The use of branching programs has excited some controversy over whether a straight line program is as effective as one which branches in response to the behavior of the learner. Skinner maintains that the important factor is the rate of learning and consequently a straight line program is usually adequate. Crowder maintains that different branches will be required for different learners. Neither position is bulwarked by a body of data, but Skinner's position is the more systematic when viewed against the perspective of a science of behavior.

The question of review is receiving some experimental attention, but the only findings presented to date offer no support for the inclusion of review items.

Recommendations.- Before authoritative rules of programming can be developed, the questions raised above must be answered. But conclusive answers can only be obtained for specified programs under specified conditions. It is not enough to ask whether prompting is more efficient than confirmation; the question should specify the type of subject matter and probably even the type of learner (young, old, bright, dull). Much of the equivocation in the studies cited in this section is due to the fact that many different kinds of miniature programs are used. What is needed is one basic or prototypical program which can be varied with respect to many variables. In that way an efficient program can ultimately be developed and sound answers can be found to questions about prompting, confirmation, explanations, size-of-step, response modes, and review items.

This writer sees little advantage to testing hypotheses about programming methods using ad hoc programs. If the ultimate purpose of teaching machine research is to develop useful programs and efficient devices, then the place to begin is with a given subject matter. Let the investigator define his task: the teaching of French, electronics, navigation, trouble shooting, calculus, psychology, or whatever he chooses. We would expect that in choosing a given subject matter, he performs a job analysis, that is, he defines the objectives of his teaching. These objectives should generate some kind of criterion test, whether it be an achievement examination or an on-the-job test. Once the subject matter is chosen, the writing of the course begins. The course (or enough of it to warrant testing) should be given and the results of the criterion tests noted. Once given it can be revised; the revisions which follow can and should incorporate the various questions about programming which have been raised in the section above. By systematically testing each of the questions, a refined program should take shape. At some point in this process, the investigator will wish to compare the teaching machine program with other techniques of teaching. If he makes these comparisons before the teaching machine program is refined, he may conclude falsely that automated teaching is inferior. However, such comparisons can be useful in shaping the final program. This problem will be discussed again in section III, below.

II. TEACHING MACHINES[1]

A number of investigators have pointed out that the construction of teaching machines is far ahead of the construction of programs (Homme and Glaser, 1959a; Gilbert, 1959). There are probably three different kinds of teaching machines for every program and if we count only the completed programs (excluding the miniature programs), the ratio of machines to programs may be as high as 20 to 1. Even the most ardent designer of hardware machines would admit that this is a serious imbalance. Some (Homme and Glaser, 1959b; Gilbert, 1959) have suggested that this imbalance results from the fact that it is more intriguing to build machines than to write programs. However, there are many questions about the design of the machines that bear important relationships to program construction and the psychology of learning. The following questions will be considered in this section: (1) Is it desirable to repeat incorrect items? (2) What display characteristics should be incorporated into teaching machines? (3) Does the teaching machine possess motivational properties? (4) Is machine pacing desirable? (5) What is the most efficient way to provide feedback or reinforcement?

Repeating Incorrect Items

One of the characteristics of modern teaching machines, particularly the early devices like Skinner's disk machine (1958), is the eliminating of items which the learner has responded to correctly. By eliminating correct items, the learner is able to repeat the incorrect items without reviewing the whole set. Although this item elimination feature is an expensive one in terms of the cost of constructing machines, it was thought to be sufficiently important to justify the expense. However, recently the question of repeating incorrect items has been reconsidered in the light of small-step programs. If the small-step technique of programming is espoused, then a good program should be very nearly free of errors. If that be the case, then there is little reason to be concerned about the repetition of incorrect items. Furthermore, Skinner[2] reports some recent research which shows that the repetition of incorrect items provided no additional benefits to learning. However, the question of item repetition may still be very relevant in programs which do not follow the principles of progressive approximations and small steps. If programs are seeded with error items, then item repetition may have a decided effect.

Display Characteristics

The typical teaching machine presents material by means of discrete visual displays. While some of the new devices have an audio component, the majority of machines are entirely visual. In the visual machine, the material is usually stored on disks, cards, or paper tapes or films, and presented in a controlled sequence by mechanical, electromechanical, or electronic means. In the early design of machines, little attention was paid to possible variations in the visual display characteristics.

---

1. Fry, Bryan, and Rigney (1959) present a summary of the types of teaching machines which have been used in the past and which are in current use. Although their list is not exhaustive, the 26 machines described do give a picture of the fast developing technology of teaching machine construction.
2. Questionnaire

Recently attention has been drawn to possible ways of varying the visual display so as to make better use of prompting techniques. A study now under way at Harvard University under the general direction of Skinner is concerned with the quantification of formal prompts.[1] A paired-associate learning paradigm involving German words and their equivalent is being used. The German word, the stimulus member of the pair, appears before the subject on an opaque screen. The English equivalent, the response member, is presented under it, well out of focus. The subject attempts to provide the English equivalent, and when he is unable to do so, he presses a switch, gradually causing the focus to improve until he is able to give the English equivalent. As learning progresses, less and less focusing will probably be required.

Silverman (1959b) found that paired-associate learning could be improved by making some of the stimulus items in the paired-associate list distinctive. While the distinctive cues were actually a part of the program, they could have been device controlled by providing different kinds of frames for some items. In this way it would be possible to vary the display characteristics of any frame within a program.

Surprisingly little work has been done with audio presentation in teaching machines. However, some research is in progress. Skinner[2] is working on an auditory supplement to his present machine. Tucker[3] reports that his group is going to investigate the use of audio techniques in self-instructional teaching. Silverman (1959a) is investigating the use of audio presentation, visual presentation, and combined audio-visual presentation. It is likely that we shall see a number of studies dealing with audio devices in the near future.

It was pointed out earlier that all teaching machines are not hardware devices. The programmed textbook described by Homme and Glaser (1959b) and Crowder's (1959) scrambled textbook are examples of teaching machines that do not depend upon mechanization but are, nevertheless, self-tutoring devices. The programmed textbooks present items in panels. A panel is read and a response called for in much the same way that it is done in the automated devices. To confirm his answer, the student turns to a specified page and then proceeds further according to instructions. The presentation technique used in these textbooks differs from the automated devices. One difference between these textbooks and the automated devices is that in the textbooks there is much less spatial constraint; adjacent frames or panels are not out of sight, so that a greater burden of selective attending is placed on the learner. In effect, he is asked to read and respond to the first item before he proceeds to the second item.

In the standard automated device the machine prevents the learner from seeing the second item until he is finished with the first item. Furthermore, in the programmed textbooks the student can "cheat," that is, he

---

1. Questionnaire
2. Questionnaire
3. Questionnaire: Tucker, J. Western Design. A Division of U. S. Industries, Santa Barbara, Calif., 1959

can turn to the answer before making his own response; the hardware devices usually prevent such looking ahead. However, if the findings of Irion and Briggs (1957), Cook and Kendler (1956), Kopstein and Roshal (1955), and Bryan and Schuster (1959) are relevant, then "cheating" by looking ahead to the answer may not defeat the purpose of self-instruction, but may actually enhance it.

Homme and Glaser (1959b) point out that if "cheating" is damaging to the learning process, it is probable that with truly effective programming techniques such cheating will tend to be eliminated. They state (p. 105):

> It is quite conceivable that the tendency to "cheat" as well as the current widespread insistence upon a "foolproof" hardware machine, represents nothing more than a vestige of traditional modes of educational control which will soon be displaced by other techniques. We may note that a necessary condition for predisposing the behavior we call cheating is that the tendency to emit the correct response is weak. Obviously this condition can be avoided in a programmed sequence. With a well constructed program in which non-reinforcement is minimized, pre-exposure of the answer is also minimized. Further, a well constructed sequence is, by definition, a program in which success on each frame depends on the successful performance in prior ones; skipping of frames is poorly reinforced and consequently weak. For these reasons, then, we have concluded that cheating against the program is not a problem of vital concern, and that extensive instrumentation to circumvent it is not required.

The Motivational Properties of Teaching Machines

Pressey (1959, p. 195) points out:

> It is a commonplace of educational research that first trials even of bad ideas usually come out well, because only teachers interested in them first try them, and pupils like the novelty and special attention.

There has been some speculation that novelty has played an important role in teaching machine training. References are frequently made to the "pinball machine effect" and the "hardware" effect, effects which may disappear in time, but which may operate in the early stages of learning to give a false picture of the effectiveness of automated teaching. Skinner (1958) has suggested that the reinforcing mechanism of his early machine may tend to produce a high level of motivation. He further suggests that this high level of motivation may cause the learner to respond too quickly and carelessly during the first cycle of the program. To counteract this effect, Skinner recommends using a slowing technique whenever an error occurs. One such technique would be to impose a five-second delay in presenting the next item, after the error occurs.

Porter (1959) found that the novelty value of teaching machines was

high when they were first introduced into class, but he saw no evidence of lessening student (elementary school) motivation after continued use. The weekly test scores for the first and second half of the lessons (spelling) were computed and no significant differences were found.

Pacing

According to Skinner, one of the key characteristics of teaching machines is that the learner is able to proceed at his own pace; he can budget his time as he desires. While, on the surface, self-pacing appears to have advantages, there has been some speculation that automatic pacing may, in some situations, produce efficient learning. These situations may be a function of the subject matter and/or individual differences. Briggs, Plashinski, and Jones (1959) compared machine pacing and self-pacing, using a 20 item paired-associate list presented on the Subject Matter Trainer. Two groups of subjects were used; one group proceeded at its own pace, while pacing was machine-controlled for the other group. No difference was found between the two groups on a test trial.

Silverman, using part of a program in basic electricity, presented on the Skinner disk machine, found no significant difference between subjects paced by external signals and subjects who were self-paced. Aside from reports about "boredom," it made little difference whether the pacing was moderate or very slow.

Feedback and Reinforcement

One of the most important features of teaching machines is the immediate feedback which they provide the learner. While there is evidence supporting the view that confirmation in the form of knowledge of results is an important factor (MacPherson, Dees, and Grindley, 1948, 1949), there is no unanimous agreement whether such confirmation is effective because it acts in concert with motivation or because it simply supplies information. Skinner and his associates consider feedback in the form of confirmation to be a major source of reinforcement. For Skinner, reinforcement refers only to those operations which strengthen learning; whether this strengthening is the result of the feedback of information per se (Tolman, 1932), whether it is the result of need reduction (Hull, 1943), or whether it is the result of protecting the associative connections (Guthrie, 1935) is not a central issue.

However, some writers (Gagne, 1958; Carr, 1959) do make a distinction between confirmation and reinforcement. Gagne reasons that when motivation depends upon the nature of the task, as when successful completion of the task is a goal, reinforcement is provided in the form of knowledge of results. Carr takes Gagne's position to mean that confirmation is reinforcing only if the learner's motivation is intrinsic to the task being learned. Thus, Carr distinguishes between confirmation and reinforcement, arguing that they are two distinct classes of variables. He cites Porter (1957) and Keislar (reference in Carr, 1959), both of whom considered the possibility that confirmation may lose its reinforcing property over a period of time because of a weakening of the "novelty effect."

The confirmation-reinforcement issue may be a pseudo issue. Those who propose that confirmation is reinforcing do not overlook the possible value of other reinforcers (Pressey, 1950; Skinner, 1954); neither do they deny the fact that certain conditions can enhance or weaken the reinforcing power of confirmation. Pseudo issue or not, an heuristic purpose may be served by attempting to keep confirmation and reinforcement separate. Perhaps the one provides information and the other maintains motivation. In certain situations knowledge of results may do both, while in other situations supplementary techniques may be required to meet the motivating demands of the situation. In tasks which command considerable learner interest, confirmation will probably serve as a powerful reinforcement. In situations where the task is not intrinsically interesting, or where the learners are under considerable pressure to achieve, praise as well as confirmation would probably be necessary to insure efficient learning. In this regard, there is probably a difference between the reinforcement needs of students taking elective courses and students taking required courses, and between armed forces trainees who volunteer for a certain type of training and those who are assigned to that training.

Summary

Conclusions.- While the construction of teaching machines continues apace, many fundamental questions about what properties a teaching machine should contain have yet to be answered satisfactorily. Answers are needed to the questions discussed in the previous section, as well as to the specific machine-directed questions raised in this section. The question of whether a machine should contain an item elimination or item repetition feature has been answered only in part. It depends upon whether or not a program of small, error-free steps is used. It has been reported that repeating incorrect items in a small-step program is not necessary. However, no answer is available for large-step programs.

Little has been done with the possible variations in the display characteristics of teaching machines. It is probable that future work on display features will await further work on the problem of prompting.

The question whether machines should be "cheat-proof" has been considered, and the tentative answer is, "no." Of course, such an answer refers only to the teaching function of the machine and not to any testing function.

The "novelty" or motivating effects of teaching machines, while attracting some interest, have received little experimental attention. The one study (Porter, 1959) bearing on this issue used a very simple device. Because of the simplicity of the device and because the subjects were elementary school children, the finding that there was no apparent lessening of motivation over a period of six months or more cannot be regarded as conclusive evidence that automated teaching devices do not produce transient motivational effects.

The question of self-pacing versus machine pacing has relevance for the design of future teaching machines. At present there is no experimental evidence to support Skinner's contention that self-pacing is more effective than machine pacing.

Where confirmation or knowledge of results is thought to be the important source of reinforcement, machines are constructed to provide immediate confirmation But questions have been raised whether confirmation is the major source of reinforcement, or whether automated instruction will require supplementary forms of reinforcement. This is an empirical question with important ramifications for education and training.

Recommendations.- There is little point in designing elaborate automated teaching devices in the absence of knowledge about programming requirements, item repetition, display features, "cheating," motivating effects, pacing, and reinforcement requirements. Beginning with the machine instead of the program is simply not efficient; it is likely to cloud our understanding of automated teaching by focusing our attention on secondary variables. Nevertheless, the teaching device is an important single variable, since it presents the program and responds to the learner. Device features and program features are not separable. A good device depends on a good program, and efficient automated teaching depends on good programs presented by means of good devices. In discussing research strategies related to teaching devices and programs, the question is one of priority rather than ultimate importance. If someone is primarily interested in constructing automated teaching devices, he should first find or construct a good program and try to improve it by experimenting with devices. In this way both the program and the device will be improved. In view of the present state of the teaching machine art, it is bound to be more fruitful to modify devices to present programs more effectively than to modify programs to fit the restrictions of a device.

III. AUTOMATED SELF-INSTRUCTION AND CONVENTIONAL INSTRUCTION

The ultimate test of the effectiveness of automated self-instruction techniques will be found in studies comparing these techniques with other teaching techniques. However, with the present state of teaching machine programs in mind, many investigators argue that comparative, control group studies are as yet premature. Those who argue this way insist that meaningful comparisons can only be made when good programs and machines are available. They insist that programs and good programming techniques must first be developed and refined before automated instruction is compared to conventional instruction. In spite of these assertions, some investigators have set up control group studies. Many of these studies are now in progress and results are not available; for some, however, results have been reported and those studies will be reviewed below.

The first systematic study of a type of automated teaching device was done by Little (1934). He compared a small multiple-choice device similar to Pressey's 1926 model (1926), an automatic test scoring device, and conventional classroom teaching. The subjects were college students and the subject matter was educational psychology. The two experimental groups (multiple-choice device and automatic test scoring device) were matched with the control groups for intelligence and pretest knowledge of the subject matter. The findings indicated that both experimental groups were superior to the control groups in a multiple-choice final examination. Little reported that differences between the multiple-choice group and the corresponding

control group were greatest for those students who achieved relatively low scores on the criterion test.

Jensen (1949) reports a study comparing 24 "superior" college students taking a special course in educational psychology with students taking the regular course. In the special course the students did not attend the regular class meetings, but an "independent study laboratory" was made available to them every afternoon. This "independent study laboratory" contained practice tests on a punch-board of the type described by Pressey (1950).[1] Textbooks and page references were available, so that students could check their incorrect responses on the practice tests. Laboratory instructors with whom the students could discuss any aspect of the course were present. When the students taking the special courses were compared with the regular students, the basis of comparison being a standard midterm and final examination, it was found that the special course students generally received higher examination grades.

In a study like that of Jensen's, it is difficult to know what factors contributed to the superior performance of the special course students. They were "superior" students; they were "encouraged" to test themselves using the punch-board device; they were "encouraged" to discuss course matters with each other and with an instructor; and they were clearly aware that they had been singled out for special attention. Any one, or some combination of two or more of these factors, could have accounted for the results. Jensen's study may have demonstrated that "superior" students need not attend formal class meetings, but the study tells us little about automated self-tutoring.

Angell and Troyer (1948) and Angell (1949) report that the use of a Pressey type punch-board on an examination resulted in learning. In one study (Angell, 1949) an experimental group used the multiple-choice punch-boards during three one-hour examinations, while the control group was given the standard multiple-choice type of examination. The course was elementary chemistry and the subjects were college freshmen. The measure of learning, the final examination scores for each group, showed the experimental group to be superior.

Pressey punch-boards were used by Briggs (1950), who attempted to demonstrate that "superior" college students could take accelerated training with a minimum of regular formal class meetings. Briggs' purpose was similar to that of Jensen (1949). His subject matter was also educational psychology, and his results were similar to Jensen's in that more of the accelerated students achieved higher grades than did the regular students. However, as far

---

1. The Pressey punch-board is a card about 3 inches by 5 inches and 1/2 inch thick. On the face the punch-board has thirty rows of four holes. A key sheet inside the card has holes opposite the right answer for the particular test being used. The student indicates his response to a test item by punching a pencil through one of the four holes for a particular item. If the hole punched is a correct answer, the pencil goes deep into the card and makes a large hole in the answer sheet. If the answer is wrong, the pencil breaks only a small hole in the paper, but does not penetrate deeply.

as teaching machines are concerned, Briggs' study suffers from many of the same shortcomings that confounded the Jensen study.

One of the most extensive teaching machine studies done to date was that done by Pressey (1950). The teaching machine used was the punch-board and the subjects were college students. In one series of experiments the subject matter was Russian vocabulary, difficult English vocabulary, and psychology. In this series of studies Pressey found that students taking a punch-board test do show some immediate learning as a result of the test and that repeating the same test is a practicable method of teaching.

In a second series of studies Pressey compared three groups of students taking a psychology course under different conditions. One group used the punch-board for practice tests; a second group took practice tests but without the punch-board. A third group had no practice tests in any form. The results indicated a slight superiority on a criterion test for the punch-board and practice test groups. When, in a third series of studies, intensive use was made of punch-board practice tests closely integrated with the subject matter, the punch-board groups did significantly better on a final examination than no-practice test groups. In a fourth series of studies three large groups of subjects were compared. One group took no-practice tests, a second group took practice tests without punch-boards, and a third group took practice tests with the punch-boards. The punch-board group scored higher on midterm examinations than either of the other two groups, and the practice-test group was superior to the no-practice group. Jones (1954) confirmed many of Pressey's findings in a series of studies performed at Ohio State University.

Porter (1959) carried out an experiment in teaching machine instruction with elementary school children in the second and seventh grades. The subject matter was spelling and the instruction took place over a twenty-two week period. Porter found that on both the second and sixth grade levels, standardized achievement tests showed the teaching machine groups to be superior to the conventional teaching groups. An interesting finding was the absence of a relationship between intelligence test scores and achievement in the experimental groups (rho = -.128), and the presence of a positive relationship in the control groups (rho = .343).

Evans, Glaser, and Homme (1959) compared verbal learning sequences of the teaching machine type with learning by means of a standard textbook. In one study a group of 17 undergraduate students in psychology was instructed to learn the material covered in ten pages of a standard statistical text. A second group of 16 students was given a "programmed text" (of the type described earlier) covering the same material. The two groups were not statistically different from each other on a multiple-choice test. However, the programmed text groups were less variable and were slightly higher in their scores than the control group.

In a second study by Evans, Glaser, and Homme, two new groups were given the same material used in the first study, but a constructed answer test was used in place of the multiple-choice test. The results were similar to those of the first study but, again, not statistically reliable. A third study used the materials from part of a course in the fundamentals of music.

In this study the programmed text group was superior to the control group on a test of musical notation. Probability values were combined for the three studies, and the result was a probability of less than .05 for the combined results, indicating that for the combination of three studies the programmed text led to better learning than did the standard text. The authors also point out that the programmed text groups took significantly more time to complete their sequences than the textbook group took to read their material.

Kormondy (1959) at Oberlin College has completed one phase of a study of the comparative effectiveness of a teaching machine program in a section of a course in introductory zoology dealing with basic genetics. A part of Kormondy's study was described above in the section dealing with review items. The complete study involved five groups of approximately 40 subjects each. Group "Alpha" used a programmed booklet (similar to Homme and Glaser's programmed textbook) comprising 461 items, each of which calls for one or more constructed responses. Confirmation of the correctness of a response was obtained by turning a page. (It should be noted that the programming techniques used by Kormondy are similar to those employed by Skinner and his associates.) The "Alpha" group did not attend the three lectures covered by the material in the programmed booklet. Group "Beta" used the same kind of programmed booklet, but in addition had 91 review items. This group, like "Alpha," did not attend the three lectures. Group "Gamma" used a textbook version of the "Alpha" program, which included some of the review items used in group "Beta." The "Gamma" group also did not attend the lectures. The textbook version differed from the programmed workbook in that the learner was not called upon to give responses. Group "Delta" attended the three lectures and was strongly encouraged to study for six hours outside of class using the regular textbook. (Kormondy reports that this was done by almost all of the subjects.) Group "Epsilon" attended the three lectures, but was not motivated other than being told that this was a control group. (Kormondy states that almost all attended the three lectures and that most reported having studied for three or four hours.)

All subjects were given a pretest and shortly after completing the program, a posttest. (Another posttest is to be given approximately six months after the learning experience.) Learning was measured by taking the differences between the posttest and pretest scores and dividing them by one-half. An analysis of covariance was performed with the difference scores adjusted in terms of pretest scores. The analysis indicated: (1) "Alpha," "Beta," and "Gamma" were not significantly different from each other, but the difference score for each was significantly higher than each "Epsilon" and higher than "Delta," but not quite significantly so. (2) "Delta's" difference score was higher than "Epsilon's," but not significantly so. (3) The "Gamma" group made the most improvement in the least amount of time (measures were taken of time spent in studying).

The Kormondy study is similar to one being done as part of a larger project at Earlham College, Richmond, Indiana, under the direction of Barlow (1959). The results of one phase of the Earlham study (Kormondy, 1960) are similar to those reported by Kormondy for the Oberlin study, in that there were no differences among the first three groups, the "Alpha," "Beta," and "Gamma" groups. However, in the Earlham study the control group (Kormondy's "Epsilon"), which was given no special instructions and simply attended the regular classes, performed as well as the first three groups. The "Delta"

group (referred to as Xi in the Earlham study) was significantly inferior to the other four groups. This was the group that used the regular textbook, but did not attend the lectures.

Thus, the Oberlin and Earlham studies fail to show any significant advantage for the programmed booklet over the textbook version of the booklet. However, both studies do tend to give some support to programming per se.

In a somewhat informal study Hatch (1959) attempted to evaluate the effectiveness of a self-tutoring approach applied to pilot training. Two matched groups of Air Force pilots were used in the study. For one group a game-like self-tutoring device originally developed by the U. S. Naval Training Devices Center was installed in the crew lounge. The device was a card machine that presented instrument flying information to the subject by means of multiple-choice items. Exposure to the device was voluntary. The pilots were briefed so as not to arouse suspicion about the true purpose of the device. They were told during a pretest (which both groups took) that the purpose of the test was to provide item analysis data of the items to be developed for the devices. No device was available in the crew lounge for the control group. After a two-month period both groups were posttested. The group that had been exposed to the device showed a slight but significant improvement over its pretest scores, while the control group showed no improvement.

Summary

Conclusions.- The control group studies reported above generally show that teaching machines can promote learning, but it would be unwarranted to conclude at this time that automated teaching is superior to conventional teaching. The only studies reported thus far with adequate control conditions are the Oberlin and Earlham studies, and Porter's study. In two of these (Oberlin and Porter) the groups taking a programmed course, teaching machine, workbook, or textbook, were generally superior to the groups exposed to conventional classroom learning and to a conventional textbook.

The other studies reported in this section make limited comparisons. The punch-board studies simply demonstrate that tests which give immediate knowledge of results do promote some learning. The studies comparing programmed textbooks to standard textbooks show that subjects who use the programmed texts achieve higher scores on criterion tests than subjects who use a standard text. However, these studies do not provide information about the differences between automated teaching and conventional teaching methods.

Recommendations.- It was suggested in the section dealing with programming that we should not be prematurely concerned with global comparisons between automated teaching and conventional teaching. This does not mean that control group-experimental group comparisons should not be made at this time; these comparisons can be made to serve as a means for improving programming and machine design. The performance of a control group can be used as a standard or baseline against which to measure the effects of different programming methods and different types of machine presentation. The question in the beginning is not simply whether the experimental or control group is superior, but how can the experimental group be improved. Such improvement can be measured better when a standard in the form of a control group is available.

Ultimately, we shall want to select the most effective method of teaching, but such selection must await the development of good programs and machines. When such programs and machines are available, the choice of a teaching method can be based on the demonstrated superiority of a given method and on such practical matters as cost, facilities, personnel, and so forth.

IV. CURRENT TRENDS AND RECOMMENDATIONS

The present state of the teaching machine art reflects the undeniable importance of programs and programming methodology. It also reflects the lack of focus frequently found in a new problem area. The questions are many; the answers are few. The questions are not often clearly stated, and the few available answers tend to be equivocal. Zeal and enthusiasm have temporarily overwhelmed skepticism and momentarily obscured some basic methodological problems. However, the observer cannot help being impressed with the vigor of the teaching machine movement and the ultimate prognosis of teaching machine research.

While much of the early research has tended to be somewhat diffuse, a set of basic issues is beginning to emerge which is bound to generate significant research studies. Some of these basic issues are: (1) What is the function of prompting, and how does it relate to confirmation and reinforcement? (2) Is the successive approximation method, involving small steps, a basic method of programming? Phrasing the question differently, are we essentially interested in shaping behavior directly as Skinner implies, or are we concerned with providing information which will indirectly manifest itself in behavior? The latter position gives rise to a corollary question: Should information be programmed so as to meet the needs of specific kinds of subjects? Other problems, like response mode, explanations, review items, item repetition, pacing display characteristics, and so forth, while important, are not fundamental issues.

The problem of prompting, confirmation, and reinforcement is a fundamental one, but it is bound to the problem of direct behavior shaping. If automated teaching (or any teaching, for that matter) consists of establishing certain behaviors directly, then means must be found to elicit the behavior and then to maintain it once it is elicited. The principle of successive approximations indicates a means for gradually shaping behavior to produce the end behavior in question. Within the general principle of successive approximation it is necessary to devise ways to elicit a given response. Prompts are used to elicit the responses and confirmation <u>qua</u> reinforcement is used to strengthen the response.

If automated teaching consists of systematically providing information, then prompts and confirmation may be considered as being special classes of information. The one class, prompts, provides information early; the other class, confirmation, provides it later.

Closely linked to the shaping behavior issue is the problem of measuring the effects of automated teaching. In order to shape a response, the

response must be identified and described. If the response can be described, a criterion measure can be specified. In other words, if we know what we want to teach, we should be able to determine whether or not the teaching is effective. In this regard the behavior-shaping position has an advantage over the information-providing position, since the latter position does not commit itself so directly to a given behavior or behaviors. But the validity of both approaches still depends upon the ultimate specification of what is to be learned. Criterion tests are necessary, whether a specified response or class of responses is being shaped or whether a subject material is being taught through the giving of information.

The question that remains is: What constitutes an adequate criterion test? Is it a test which indirectly measures the learner's accomplishment, or is it a test which provides a direct measure of the behavior in question? These questions introduce the problem of specifying what is to be taught. A thorough job analysis of the task to be learned should permit the programmer qua teacher to specify precisely what responses he is trying to establish and under what stimulus control. A good job analysis immediately generates criterion measures; whether these measures take the form of a paper and pencil test or an on-the-job test depends on the subject material.

One of the major shortcomings in present-day teaching machine research is the lack of precise job analyses. This failure is more apparent in education where complete courses are given, than in training situations where tasks are more circumscribed. In training situations the job is usually the criterion, and the job analyses are a matter of course. But where the aims of teaching are not concretely defined, there is a tendency to ignore the job analysis process.

Where job analyses are done, there is still the problem of transfer of training to be considered, since some types of training and education are directed toward a variety of tasks. It has been suggested (Kendler, 1959) that there are differences in the ability to transfer learning, depending upon the kind of learning, kind of task, and kind of learner. This problem should be given serious consideration in job analyses, in constructing criterion tests, and in ultimately constructing teaching machine programs.

Much attention has been paid to the problem of developing specific rules for programming subject matter, but little attention has been paid to setting up equally specific rules for analyzing the to-be-learned material. It is often implicitly assumed that the job analysis of the subject matter is done in the course of programming. However, it is becoming increasingly apparent that job analyses must be more explicit and systematic. Rules and principles are needed at the criterion end as well as at the programming end. A possible model to follow is that developed in the Department of Engineering at the University of California at Los Angeles (1958) under the direction of a Curriculum Study Committee now headed by A. B. Rosenstein. This committee has prepared a set of rules and principles to identify the major elements which make up the content of engineering courses.

Once the definition of what is to be learned is established, there is still the question of how to measure performance. In most of the recent teaching machine studies the degree of learning is measured by some type of

criterion test. Subjects are distinguished on the basis of the number of items which they answer correctly. However, one of the strong positions in automated teaching, that of B. F. Skinner, is based on the contention that the basic criterion should be the amount of material learned, or, more precisely, how much of the course the learner has completed. While some criterion tests might measure this, most such tests are confounded by individual differences in intelligence or performance skills. According to Skinner, two subjects completing the same program should "know" an equal amount, although it might take one subject half as long to complete the program. This suggests that for Skinner the best criterion is the amount of subject matter completed. Whether this is a sound criterion remains to be demonstrated in the face of job analyses, transfer of training tests, and long-term retention tests.

Questions about individual differences are generally still unanswered. While some investigators allude to the problem, the majority has chosen to postpone considerations of individual differences. Whether such questions can be postponed depends on the aims and, to some extent, the philosophy of the programmer. Individual differences are considered by the intrinsic programming group which sets up program branches to meet the needs of different subjects. Individual differences tend to be ignored by Skinner, who recommends using programs which can ultimately be mastered by most individuals. The extent of mastery in a given amount of time will depend upon the individual, but all subjects who eventually complete a program will be equal with respect to the material covered in that program. In this regard it is interesting to note that Porter (1959) found no relationship between intelligence test scores and achievement in a group of children trained with a teaching machine, but he found a significant positive relationship in a group trained by conventional teaching methods.

If Porter's findings can be extended to a variety of adult subjects, the programmer could afford to ignore individual differences. In the absence of such findings, the present-day programmer cannot ignore individual differences if he is programming with a practical end in mind. If the programming is aimed at developing an efficient course in electronic trouble shooting, or a course in Binary numbers for Naval technicians, or any pragmatic course, then it would be useful to include individual difference dimensions in developing and testing the program. These dimensions might be based on prior training, intelligence test scores, aptitude test scores, or personality factors.

REFERENCES

Angell, G. W. Effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. J. Educ. Res., 1949, 42, 391-394

Angell, G. W., and Troyer, M. E. A new self scoring device for improving instruction. School and Society, 1948, 67, 84-85

Barlow, J. A. New Instruction Media: Self instruction, guided instruction, and the role of the teacher. Project Description: HEW Project No. 7-12-026.00, 1959

Beck, J. On some methods of programming. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959, pp. 55-62

Briggs, L. J. The development and appraisal of special procedures for superior students and an analysis of the effects of "knowledge of results." In Ohio State Univ., Abstracts of Dissertations, 1948-1949. Columbus, O., 1950, 58, 41-49

Briggs, L. J. Two more self instructional devices. Psychol. Rep., 1958, 4, 671-676

Briggs, L. J., Plashinski, D., and Jones, D. L. Self pacing versus automatic pacing of practice on the subject matter trainer. Unpublished paper, cited in Darby, C. L. An annotated bibliography on the automation of instruction. U. S. Army Air Defense, Human Research Unit, Fort Bliss, Texas, April 1959

Bryan, G. L., and Rigney, J. W. An evaluation of a method for shipboard training in operations knowledge. Tech. Rep. No. 18, Personnel and Training Branch, Psychological Services Division, ONR, 1956

Bryan, G. L., and Schuster, D. H. The effectiveness of guidance and explanations in trouble shooting training. Tech. Rep. No. 28. Electronics Personnel Research Group, University of Southern California, 1959

Carr, W. J. Self instructional devices: A review of current concepts. Tech. Rep., TR 59-503, Wright Air Development Center, 1959

Cook, J. O. Supplementary Report: Processes underlying learning a single paired-associate item. J. exp. Psychol., 1958, 56, 455

Cook, J. O., and Kendler, Tracy S. A theoretical model to explain some paired-associate data. In G. Finch and F. Cameron (Eds.) Symposium on Air Force Human Engineering, Personnel and Training Research. Washington, D. C.: National Academy of Sciences - National Research Council, Publ. No. 455, 1956, pp. 90-98

Coulson, J. E., and Silberman, H. F. Results of initial experiments in automated teaching. Rep. No. SP-73, Systems Development Corporation, Santa Monica, Calif., July 1959

Crowder, N. A. Automatic tutoring by means of intrinsic programming. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959, pp. 109-116

Curriculum Study Committee. Second Progress Report. Department of Engineering, University of California, Los Angeles, May 1958

Evans, J. L., Glaser, R., and Homme, L. E. The effects of variations in the properties of verbal learning sequences of the "teaching machine type." Paper read at Eastern Psychological Association, Atlantic City, April 1959

Ferster, C. B. The role of review material in continuous programming with teaching machines. Proposal submitted to Commissioner of Education, U. S. Office of Education, Dept. of Health, Education and Welfare, June 1959

Fry, E. B. Teaching machines: An investigation of constructed versus multiple choice response modes. Unpublished doctoral dissertation, University of Southern California, Los Angeles, 1960

Fry, E. B., Bryan, G. L., and Rigney, J. W. Teaching machines: An annotated bibliography. Tech. Rep. No. 28, Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, November 1959

Gagne, R. M. A Review of Factors in Learning Efficiency. AFOSR TN58-924. Air Force Office of Scientific Research. Air Research and Development Command. U. S. Air Force. Washington, D. C., November 1958

Gilbert, T. F. Some issues involved in selecting an automatic teaching device. A report to Bell Telephone Laboratories, Inc., Murray Hill, N. J., September 1958(a)

Gilbert, T. F. An early approximation of principles of analyzing and revising self-instructional programs. Paper read at the Conference on Teaching Machines, University of Pennsylvania, December 1958(b)

Gilbert, T. F. On the relevance of laboratory investigations of learning to self instructional programming. Paper presented at the American Psychological Association Meeting, Cincinnati, O., September 1959

Guthrie, E. R. The Psychology of Learning. New York: Harper, 1935

Hatch, R. S. An evaluation of the effectiveness of a self tutoring approach applied to pilot training. Tech. Rep. TR 59-320, Wright Air Development Center, 1959

Holland, J. G. Teaching machines: an application of principles from the laboratory. Paper read at the 1959 Invitational Conference of Testing Problems sponsored by the Educational Testing Service, December 1959

Homme, L., and Glaser, R. Problems in programming verbal learning sequences. Paper read at the Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Annual Convention, Cincinnati, O., September 1959(a)

Homme, L., and Glaser, R. Relationships between the programmed textbook and teaching machines. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959(b), pp. 103-108

Hull, C. L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943

Irion, A. L., and Briggs, L. J. Learning task and mode of operation variables in the use of the subject matter trainer. Tech. Rep. AFPTRC TR 57-8, Air Force Personnel and Training Research Center, Lackland AFB, Texas, October 1957

Jensen, B. T. An independent study laboratory using self scoring tests. J. Educ. Res., 1949, 43, 134-137

Jones, R. S. Integration of instructional with self scoring measuring procedures. In Ohio State University, Abstracts of Dissertations, Columbus, O., 1954

Kendler, H. H. Teaching machines and psychological theory. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959, pp. 177-185

Kitson, H. D. A study of the output of workers under a particular wage incentive. Univ. J. Business, 1922, 1, 54-68 (Hawthorne Study)

Kopstein, F. F., and Roshal, S. M. Methods of presenting word pairs as a factor in foreign vocabulary learning. Amer. Psychol., 1955, 10, 354 (Abstract)

Kormondy, E. J. Oberlin Experiment in Self-Instruction in General Zoology: Preliminary Analyses. Oberlin College, Oberlin, O., December 1959

Kormondy, E. J. Earlham Experiment in Self Instruction in General Zoology: Preliminary Analyses. Oberlin, O., February 1960

Little, J. K. Results of use of machines for testing and for drill, upon learning in educational psychology. J. exp. Educ. 1934, 3, 45-49

Lumsdaine, A. A. Optimal cueing methods applicable to response-prompting for automated-instruction. Description of AFOSR Project, American Institute for Research, Pittsburgh, Penn., June 1959(a)

Lumsdaine, A. A. Teaching machines and self-instructional materials. Audio-Visual Communication Rev., 7, 163-181, 1959(b)

MacPherson, S. J., Dees, V., and Grindley, G. C. The effect of knowledge of results on learning and performance. II, III. Quart. J. exp. Psychol., 1948, 1949, 1, 68-78, 167-174

Porter, D. A critical review of a portion of the literature on teaching devices. Harv. Educ. Rev., 1957, 27, 126-147

Porter, D. Teaching machines. Harv. Grad. Sch. Educ. Assoc. Bull., 1958, 3, 2-5

Porter, D. Some effects of year long teaching machine instruction. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959, pp. 85-90

Pressey, S. L. A simple device for teaching, testing, and research in learning. School and Society, 1926, 23, 373-376

Pressey, S. L. Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. J. of Psychol., 1950, 29, 417-447

Pressey, S. L. Certain major psycho-educational issues appearing in the conference on teaching machines. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959, pp. 187-198

Sidowski, J. B., Kopstein, F. F., and Shillestead, Isabel J. Proposal for an Experiment to Examine the Influence of Several Factors on the Efficacy of a Prompting versus a Confirmation Procedure in Programming. 59-1, Wright Air Development Center, August 1959

Silverman, R. E. Study Plan for Project 0-4204. Presentation of Information on Teaching Machines. Human Engineering Contract No. N61339-507. U. S. Naval Training Devices Center, 1959(a)

Silverman, R. E. The Use of Context Cues in Teaching Machines. Human Engineering Contract No. N61339-507. U. S. Naval Training Devices Center, 1959(b)

Skinner, B. F. The science of learning and the art of teaching. Harv. Educ. Rev., 1954, 24, 86-97

Skinner, B. F. Teaching machines. Science, 1958, 128, 969-977

Skinner, B. F., and Holland, J. G. Principal Features of a Useful Teaching Machine. Final Report to Ford Foundation, 1958

Smith, D. E. P. Speculations: Characteristics of successful programs and programmers. In E. Galanter (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley & Sons, 1959, pp. 91-102

Tolman, E. C. Purposive Behavior in Animals and Men. New York: Appleton-Century-Crofts, 1932

Zuckerman, Carl. Study Plan for Research in the Automation of Teaching. U. S. Naval Training Devices Center. Human Engineering Contract No. N61339-661, 1959

TECHNICAL APPENDIX

Appendix A

Questionnaire Sent to Persons Doing Research
in the Area of Automated Teaching

NEW YORK UNIVERSITY

University College of Arts and Science

University Heights, New York 53, N. Y.

October 20, 1959

Dear

I have been asked by the United States Naval Training Device Center in Port Washington, New York, to gather information about current research on automated teaching devices. May I enlist your cooperation by asking you to respond to the following questions? I have left blank spaces for your answer and have enclosed a self-addressed stamped envelope.

1. What is the general area of your research? Are you primarily concerned with the devices, with programs, with programming techniques, with the place of teaching machines in education, etc.?

2. What specific problems are you investigating?

3. What is the general nature of your approach (e.g., purely empirical, some particular learning theory, information theory, etc.)?

4. What contributions do you think your findings will make (e.g., the design of better machines, the development of an efficient program, etc.)?

5. Who supports your research?

   a. How many professional people work on the research project?

6. Can you suggest any other research programs that are similar to or closely allied to yours?

I appreciate your cooperation. Please do not hesitate to send any additional information (in the form of letters, reports, or reprints) that you think is relevant.

Cordially,

Robert E. Silverman
Associate Professor of Psychology

Reproduction in whole or in part is permitted
for any purpose of the United States Government.

DISTRIBUTION LIST

SNDL

A2A	(10)
A3	(7)
A5	(14)
A6	(3)
B1	(5)
E4	(6)
E5	(4)
E16	(6)
G9A	(1)
H5	(2)
H8	(2)
H24	(1)
J28	(2)
J31	(1)
J33	(1)
J37	(2)
J47	(2)
J60	(1)
J73	(1)
L7	(3)
L13	(1)
R20	(1)
W1B	(1)
W3A	(1)
W3B	(1)
W3C	(1)
W3F	(1)
W4H	(1)
W7C	(3)
W7E	(1)
W7X	(1)
W10A	(1)
24G	(1)
28K	(1)
42A	(1)
E2D	(1)
E6D	(9)

AIR FORCE

Department of AF Library, Air University, Maxwell AFB, Montgomery, Alabama

CDR AF Cambridge Research Center, Air Research & Development Command, (Attn: CROS), L. G. Hanscom Field, Bedford, Massachusetts

HQ Air Research & Development Command, Life Sciences Directorate (RDSPC-LS) (Attn: MAJ John L. Masin), Andrews AFB, Washington 25, D. C.

HQ USAF School of Aviation Medicine, Library Services Section, Randolph AFB, San Antonio, Texas

CDR WADD (Attn: Dr. M. Rockway), Psychology Branch, WPAFB Ohio (4)

CDR WADD (Attn: WCLDPTR) WPAFB Ohio

CDR WADD (Attn: Mr. Melvin J. Warrick) WPAFB Ohio

CDR WADD (Attn: WCRDP-5) WPAFB Ohio

ARMY

CG US Army Infantry Human Research Unit (Attn: Librarian) USCONARC, Fort Benning, Georgia

CG Aberdeen Proving Ground (Attn: Human Engineering Lab) Aberdeen, Maryland

HQ Continental Army Command G-3 Section (D&R) (Attn: LTCOL S. H. Burrow), Fort Monroe, Virginia

CIVILIAN

American Institute for Research (Attn: Dr. Arthur A. Lumsdaine), Pittsburgh Pennsylvania

Arizona State College, (Attn: Dr. Richard E. Schutz), Tempe, Arizona

Askania Regulator Company (Attn: Mr. H. Ziebolz), Chicago, Illinois

AVCO Manufacturing Corp. Human Engineering Section (Attn: Mr. Frederick W. Trabold, Jr.), Cincinnati, Ohio

Bell Telephone Laboratories (Attn: Dr. E. Z. Rothkopf), Murray Hill, New Jersey

Bendix Systems Division (Attn: Mr. John J. Cahill), Ann Arbor, Michigan

Bolt, Beranek, Newman, Inc. (Attn: Dr. J.C.R. Licklider), Cambridge, Massachusetts

Brooklyn College Department of Psychology (Attn: Dr. Carl Zuckerman), Brooklyn, N. Y.

University of California Department of Education (Attn: Mr. Evan Keislar, Los Angeles, Calif.

University of Californa, Santa Barbara College (Attn: Chancellor Samuel B. Gould), Goleta, Calif.

University of California Department of Psychology, Santa Barbara College (Attn: Dr. Robert M. Gottsdanker), Goleta, Calif.

Cambridge University Applied Psychology Unit, Cambridge, England

Columbia University Department of Industrial Engineering, (Attn: Mr. Peter Frank), New York

Robert E. Corrigan & Associates (Attn: Dr. R. E. Corrigan) Garden Grove, Calif.

Defence Research Board Applied Psychology Unit (Attn: Dr. R. B. Bromiley), Toronto, Ontario, Canada

Denison University Department of Psychology (Attn: Dr. Robert S. Carter), Granville, Ohio

Dunlap and Associates, 429 Atlantic Street, Stamford, Connecticut

Earlham College Psychology Department (Attn: Dr. John A. Barlow), Richmond, Indiana

Electric Boat Company Human Factors Section (Attn: Dr. Raymond C. Sidorsky), Groton, Connecticut

Emory University Department of Psychology (Attn: Dr. Benjamin Wyckoff), Atlanta, Georgia

Franklin Institute (Attn: Dr. Ezra Krendel), Philadelphia, Pa.

Fund for the Advancement of Education (Attn: Mr. Ronald Gross) New York, New York

General Electric Company Advanced Electronics Center (Attn: Dr. William B. Knowles), Ithaca, New York

University of Georgia Department of Psychology (Attn: Mr. Thomas F. Gilbert), Athens, Georgia

Grumman Aircraft Corporation (Attn: Mr. John D. Vandenberg), Bethpage, New York

Harvard University Psychology Laboratories, (Attn: Mr. James G. Holland) Cambridge, Massachusetts

Harvard University School of Education (Attn: Mr. Douglas Porter) Cambridge, Massachusetts

Harvard University (Attn: Professor B. F. Skinner, Cambridge, Massachusetts

HRB-Singer, Inc. Human Factors Division (Attn: Dr. George Grant), Science Park, State College, Pennsylvania

Hughes Aircraft Corp. (Attn: Dr. L. J. Briggs) Culver City, California

University of Illinois Department of Psychology, Urbana, Illinois

University of Illinois Laboratory of Aviation Psychology (Attn: Dr. Jack A. Adams), University Airport, Savoy, Illinois

Indiana University Medical Center, Institute for Psychiatric Research, (Attn: Dr. Charles B. Ferster), Indianapolis

Indiana University Department of Psychology (Attn: Dr. Irving Saltzman), Bloomington, Indiana

State University of Iowa Department of Psychology (Attn: Mr. Judson Brown), Iowa City, Iowa

State University of Iowa Department of Psychology (Attn: Dr. D. Lewis), Iowa City, Iowa

The Johns Hopkins University Department of Psychology (Attn: Dr. A. Chapanis), Baltimore, Maryland

The Johns Hopkins University Applied Physics Laboratory (Attn: Dr. J. W. Gebhard), Baltimore, Maryland

Kent State University Department of Psychology (Attn: Mrs. Idabelle K. Hoose), Kent, Ohio

Link Aviation Company (Attn: Mr. W. Wood) Binghamton, New York

Lockheed Aircraft Corporation Missile Division (Attn: Dr. J. Wissel), Sunnyvale, Calif.

Loyola University Reading Clinic (Attn: Mr. Edward B. Fry), Los Angeles, Calif.

Manhattan State Hospital (Attn: Dr. Barbara Goldenberg) Wards Island, New York, New York

University of Maryland Department of Psychology (Attn: Dr. T.G. Andrews), College Park, Md.

Medical Research Laboratory Psychology Department, Fort Knox, Kentucky

University of Michigan Department of Psychology (Attn: Dr. Paul M. Fitts), Ann Arbor, Michigan

University of Michigan Elementary School 524 (Attn: Dr. Donald E. P. Smith), Ann Arbor, Michigan

Minneapolis-Honeywell Regulator Co. Aeronautical Division (Attn: Mr. J. W. Senders), Minneapolis, Minnesota

MIT Servomechanism Laboratory (Attn: Prof. George Newton), Cambridge, Massachusetts

MIT Psychology Section, Cambridge, Massachusetts

Montana State University Department of Psychology (Attn: Dr. Carol H. Ammons), Missoula, Montana

North American Aviation Corporation (Attn: Mr. Benjamin Clymer) Columbus, Ohio

North Carolina State College Department Of Psychology (Attn: Dr. John O. Cook), Raleigh, N. C.

Nortronics Human Factors Division (Attn: Dr. Kenneth R. Colson) Anaheim, Calif.

NYU College of Engineering (Attn: Research Division), University Hgts, New York, New York

NYU College of Arts and Science, Psychology Department (Attn: Dr. Charles N. Cofer), Washington Square, New York

NYU College of Arts and Science, Psychology Department (Attn: Dr. Lawrence Karlin), Washington Square, New York

NYU College of Engineering, Psychology Department (Attn: Prof. Howard H. Kendler), Bronx 53, New York

NYU (Attn: Prof. Glen Heathers), Main 520 100 Washington Square East, New York

NYU (Attn: Mr. E. Arthur Hungerford, Jr.), Main 910, 100 Washington Square East, New York

NYU (Attn: Dean Thomas C. Pollock), Main 910, 100 Washington Square East, New York

NYU College of Arts and Science, Psychology Department (Attn: Dr. Robert S. Silverman), Washington Square, New York (10)

Oberlin College Department of Zoology, (Attn: Mr. E. J. Kormondy), Oberlin, Ohio

Ohio State University Research Foundation, Laboratory of Aviation Psychology (Attn: Dr. George E. Briggs), Columbus, Ohio

University of Pennsylvania Psychology Department (Attn: Dr. Jacob Beck), Philadelphia, Pennsylvania

Pennsylvania State University Department of Psychology (Attn: Dr. John F. Corso), University Park, Pennsylvania

University of Pittsburgh Department of Psychology (Attn: Dr. Robert Glaser), Pittsburgh, Pennsylvania

Psychological Research Associates, Inc. (Attn: Dr. Harry Older), 507 18th Street S., Arlington, Virginia

Princeton University Department of Psychology (Attn: Dr. Robert M. Gagne), Princeton, New Jersey

Radio Corporation of America, Moorestown, New Jersey

Radio Corporation of America Missile Electronics & Controls Division (Attn: Dr. Edward Gavurin) Burlington, Massachusetts
University of Rochester Department of Psychology (Attn: Dr. S.D.S. Spragg), Rochester, New York
Rochester Institute of Technology (Attn: Mr. Maurice Kessman) Rochester, New York
Rowland and Company (Attn: Dr. G. E. Rowland), 1610 Ravenswood Way, Haddonfield, New Jersey
San Diego State College Psychology Department (Attn: Dr. Joseph B. Sidowski), San Diego, Calif.
Swartzlander and Heyer (Attn: Dr. James M. Tanner), Birmingham, Alabama
Systems Development Corporation (Attn: Dr. W. Biel), Santa Monica, California (3)
Teaching Machines, Inc. (Attn: Dr. Lloyd E. Homme), 235 San Pedro, N.E., Albuquerque, N.M.
Temple University (Attn: Mr. William J. Carr), Psychology Department
Tufts University (Attn: Dr. Leonard C. Mead), Medford, Massachusetts
Tulane University (Attn: Mr. Arthur L. Irion), New Orleans, Louisiana
Utah State University Bureau of Educational Research (Attn: Mr. John R. Cragun), Logan, Utah
Western Design Division of U. S. Industries, Inc. (Attn: Mr. N.A. Crowder), Goleta, Calif. (2)
University of Wisconsin Department of Psychology (Attn: Dr. David A. Grant), Madison, Wisconsin
UCLA School of Education (Attn: Dr. Sidney L. Pressey), Los Angeles, Calif.